

THESE DE DOCTORAT DE

L'ÉCOLE NATIONALE SUPÉRIEURE MINES-TELECOM ATLANTIQUE
BRETAGNE PAYS DE LA LOIRE - IMT ATLANTIQUE

ÉCOLE DOCTORALE N° 601
*Mathématiques et Sciences et Technologies
de l'Information et de la Communication*
Spécialité : *Informatique*

Par

Olivier JACQ

Real-time detection, contextual analysis and visualisation of cyber-attacks: elaboration of the *Maritime Cyber Situational Awareness*

Thèse présentée et soutenue à Brest, le 11 janvier 2021
Unité de recherche : Lab-STICC – UMR CNRS 3285
Thèse N° : 2021IMTA0228

Rapporteurs avant soutenance :

Nora Boulahia-Cuppens
Joaquín García-Alfaro

Professeur, École Polytechnique de Montréal
Professeur, Institut Mines-Télécom et Institut Polytechnique de Paris

Composition du Jury :

Président :	Joaquín García-Alfaro	Professeur, Institut Mines-Télécom et Institut Polytechnique de Paris
Examineurs :	Nora Boulahia-Cuppens	Professeur, École Polytechnique de Montréal
	David Brosset	Maître de conférences, Arts et Métiers Sciences et Technologies
	Jacques Simonin	Directeur d'études, Institut Mines-Télécom Atlantique Bretagne Pays de la Loire
	Patrick Hébrard	Responsable recherche cyber et innovation, Naval Group
Dir. de thèse :	Yvon Kermarrec	Professeur, Institut Mines-Télécom Atlantique Bretagne Pays de la Loire

Invité

Laurent Aufrechter	Ingénieur cybersécurité, Thales
--------------------	---------------------------------

Contents

Contents	i
List of Figures	v
List of Tables	xi
I Introduction	3
II On maritime cybersecurity	11
II.1 The strategic importance of the maritime sector in France	11
II.1.1 The maritime world	11
II.1.2 The French civil maritime sector	15
II.1.3 The French military maritime sector	16
II.2 The digital transformation of the maritime world	17
II.3 The "marétique"	19
II.3.1 Onboard maritime information systems	19
II.3.2 Ports and offshore installations information systems	21
II.3.3 The peculiarities of MISs	21
II.3.4 The future of MISs	23
II.4 Cyber security of Maritime Information Systems	24
II.4.1 Cyber security and maritime environment: definitions	24
II.4.2 Inventory of MISs vulnerabilities	27

II.4.3	Exploitation of MISs vulnerabilities for offensive purposes	29
II.4.4	Sources of threats, cyber risks for MISs	31
II.5	Conclusion	31
III	<i>Maritime Cyber Situational Awareness: definition, model and contribution to maritime cyber security</i>	33
III.1	Cyber monitoring of MISs: the problem	33
III.1.1	Contributions of cyber monitoring to the difficulties of the maritime world	34
III.1.2	Inadequacy of classical cyber monitoring methods	36
III.2	Situational awareness	38
III.2.1	General context	38
III.2.2	First application of the SA in an aviation context	41
III.2.3	Application of SA to the maritime context	43
III.2.4	Operational processes for <i>Cyber Situational Awareness</i>	46
III.3	Modeling the <i>Maritime Cyber Situational Awareness</i>	51
III.3.1	Example of state of knowledge for a proper MCSA	51
III.3.2	Cyberattack modeling: the example of the ATT&CK model from MITRE	51
III.3.3	MCSA abstraction levels	54
III.3.4	MIS cyber monitoring system data flow aligned with SA operational processes	56
III.4	Benefits of visualisation for the KU of MISs and MCSA	64
III.5	Conclusion	65
IV	<i>Maritime Cyber Situational Awareness: architecture, experimentation and results</i>	67
IV.1	Design and experimentation of a platform prototype	67
IV.1.1	Onboard component	68

IV.1.2 "Shore" Component	73
IV.1.3 Software architecture for MCSA elaboration on the Naval Cyber Range	75
IV.2 Quality analysis - results	76
IV.2.1 Quality criterion Q1 - Trustworthiness	78
IV.2.2 Quality criterion Q2 - <i>Truthfulness</i>	80
IV.2.3 Quality criterion Q3 - <i>Completeness</i>	81
IV.2.4 Quality criterion Q4 - <i>Freshness</i> and <i>timeliness</i>	87
IV.2.5 Quality analysis - Conclusion	90
IV.3 Visualisation and mapping - results	91
IV.3.1 Dynamic risk assessment	96
IV.4 Application of the MCSA by domain: detection and visualisation on NMEA 0183 data, contribution to the implementation of anomaly detection by ma- chine learning	100
IV.4.1 Context	100
IV.4.2 Description of the NMEA 0183 standard	101
IV.4.3 Vulnerabilities of the NMEA 0183 standard	102
IV.4.4 Peculiarities of the NMEA 0183 standard for anomaly detection . . .	103
IV.4.5 Results	109
IV.5 Conclusion	114
V General conclusion and perspectives	117
V.1 Problematic	117
V.2 Results	118
V.3 Discussion	119
V.4 Perspectives	121
V.4.1 Data	121
V.4.2 Model	123
V.4.3 Architecture	124

Bibliography	129
Appendices	139
A <i>Information Technology and Operational Technology</i>	141
B Generic cyber risks pour maritime information systems	145
C Attack modeling on an ECDIS system using MITRE frameworks.	149
D Visualisation of a NMEA network determinism based on a talker id-based network traffic capture.	153

List of Figures

I.1	Research fields of the present thesis (source: personal work).	9
I.2	Description of the relations between the chapters of the manuscript of the present thesis (source: own work).	10
II.1	Evolution of the number (left) and of the gross tonnage (right) of ships on the 1993-2019 period (source : <i>Lloyd's Register Foundation</i> , © Lloyd's Register Foundation 2018 . All rights reserved).	12
II.2	Relative dimensions of modern ships (source: CMA/CGM, infography: Le Figaro , © Service Infographie / lefigaro.fr / 09/06/2018, reproduced with their kind authorization).	13
II.3	Strategic maritime routes for France (source: Étude "Impacts sur l'économie française de la fermeture d'un ou plusieurs détroits maritimes majeurs" of the Compagnie européenne d'intelligence stratégique, Janvier 2016. © CEIS, All rights reserved, reproduced with their kind authorization).	14
II.4	The bridge of the research vessel <i>Sikuliaq</i> (source: Wikimedia Commons , unmodified picture, author: NorthBySouthBaranof, license: CC BY-SA 4.0). . .	21
II.5	Peculiarities of MISs (source: personal work).	23
II.6	Functional, application and technical views of an information system (source: personal work).	25
II.7	Systemic analysis of cyber vulnerabilities in the maritime environment induced by digital transformation (source: personal work).	29
III.1	Benefits of cyber monitoring for the maritime sector (source: personal work).	35

III.2	Limitations of the use of classic SIEMs in the maritime context and possible solutions (source: personal work).	37
III.3	A man, as an agent, in his environment and facing a situation (source: personal work).	39
III.4	Situational Awareness elaboration modeling according to Endsley [End95], adapted with the Situation Resolution component suggested by Mc Guinness & Foy [McG00] (source: personal work).	41
III.5	Head Up Display of a Bombardier CRJ-900 passenger jet in long final at Munich airport (source: Wikimedia Commons , unmodified picture, author: Joschiki, license: CC BY-SA 3.0).	43
III.6	US Coast Guards VTS operator (source: Wikimedia Commons , unmodified picture, author: US Coast Guard, license: public domain).	45
III.7	Example of an ECDIS Human Machine Interface (source : Wikimedia Commons , unmodified picture, author: Hervé Cozanet , license : CC BY-SA 3.0).	46
III.8	Constituent processes of the CSA, work adapted from Barford and Mc Guinness & Foy [McG00] and taking into account the remarks of Franke and Brynielsson [Fra14] (source: personal work).	49
III.9	Main research subjects related to CSA, according to Peng Liu [Bar10] (source: personal work).	49
III.10	Example of a hypergraph (source: Wikimedia Commons , unmodified picture, author: Kilom691 , license: CC BY-SA 3.0).	56
III.11	MCSA modeling using super networks (source: personal work).	57
IV.1	Overview of three loops of the experimentation platform (industrial systems part) (source: personal work).	68
IV.2	Simulation environment of the navigation bridge (source: Naval cyber defense chair , reproduced with the kind permission of the author).	69
IV.3	GPS receiver (left), NMEA 2000 bus (center) and AIS receiver (right) (source: personal work).	70
IV.4	Overview of the mobility loop (the NIDS sensor is at the upper right angle of the rack) (source: Naval cyber defense chair , reproduced with the kind permission of the author).	71

IV.5	Technical modeling of the prototype with the "bridge", "communication" and "industrial" subsystems (source: personal work).	71
IV.6	Simplified technical modeling of the interconnections of the experimentation platform with the cyber prototype subsystem (source: personal work).	72
IV.7	Software architecture of the shore component for the maritime cyber monitoring process (source: personal work).	74
IV.8	End-to-end analysis of a network flow (source: personal work).	75
IV.9	End to end analysis of the logs flow (source: personal work).	76
IV.10	Possible parallel between data, information, knowledge, intelligence and situational awareness (source: personal work).	77
IV.11	Quality measurement: overview of the dimensions of quality, links with data and risks (source: personal work).	77
IV.12	Average delay of an ICMP <i>ping</i> on the WAN of the <i>Naval Cyber Range</i> platform (source: personal work).	83
IV.13	Diagram of the Round Trip Time of a satellite link to a ship at sea (source: personal work).	83
IV.14	Adapted ICMP <i>ping</i> delay on the WAN link of the platform (source: personal work).	84
IV.15	Evolution of the <i>stats.capture.kernel_packets</i> parameter for the four industrial loops (source: personal work).	85
IV.16	Evolution of the <i>stats.capture.kernel_packets</i> parameter of an IT network (source: personal work).	86
IV.17	Deterioration of the sub-criterion Q3.3 following a hardware failure simulation (source: personal work).	87
IV.18	Calculation formula of the freshness criterion Q4 (source: personal work). . .	88
IV.19	Positive evolution of the data freshness criterion during the NTP synchronization phase (source: personal work).	88
IV.20	Data timeliness following a satellite outage simulation (source: personal work). .	89
IV.21	Recovery of the delay in the counter reception (source: personal work).	90

IV.22 Multisensors data fusion (NIDS, HIDS) for anomaly detection on a MIS (source: personal work).	93
IV.23 Overview of the collected data with the flow capacity activated on a NIDS (source: personal work)	94
IV.24 Overview of data collected from a NIDS at the application level (source: personal work)	95
IV.25 Graph of the <i>dest_port</i> metadata (ordinate) in time (source: personal work).	96
IV.26 Realtime parallel coordinates graph without the proposed fix (source: personal work).	97
IV.27 Realtime parallel coordinates graph with the proposed fix (source: personal work).	98
IV.28 Use case of abnormal connexions from an ECDIS (source: personal work).	99
IV.29 visualisation and near real-time link of metadata related to an attack by the use of graphs (source: personal work).	99
IV.30 Added-value of the graph database to establish the real-time link between the signals of an attack, KU data and a potentially exploited CVE (source: personal work).	99
IV.31 Added-value of a graph database from <i>situation perception</i> to <i>situation resolution</i> (source: personal work).	100
IV.32 GPGBGA frame sample on a NMEA 0183 network (source: Wikipédia , license: CC BY-SA 3.0).	101
IV.33 The three possible levels of anomaly detection on a NMEA 0183 network, once multiplexed and encapsulated (source: personal work).	103
IV.34 Repartition, in quantity, of the type of sentences IDs on the files <i>cap1.pcapng</i> , <i>cap2.pcapng</i> and <i>cap3.pcapng</i> (source: personal work).	105
IV.35 Model of the NMEA network, talkers and sentences IDs from capture files (source: personal work).	108

IV.36	Model of the overall NMEA architecture from files <i>cap1</i> , <i>cap2</i> and <i>cap3</i> (sources: ECDIS: Wikimedia Commons , unmodified picture, author: Hervé Cozanet , license: CC BY-SA 3.0 , RADAR: Wikimedia Commons , unmodified picture, author: Clipper , license: CC BY-SA 2.5 , GPS: Wikimedia Commons , unmodified picture, author: Fairley , license: CC BY-SA 2.0).	109
IV.37	Information sharing and enhancement of the Maritime Cyber Situational Awareness using a control and alert panel on an industrial system of the Naval Cyber Range (source: personal work).	110
IV.38	Exploitation of an intrinsic vulnerability of the NMEA protocol and alerting scheme to the ECS to display the appropriate CoA (source: Naval Cyber Defence Chair , reproduced with the kind permission of the author).	112
IV.39	Exploitation of an extrinsic vulnerability of the GPS network and alerting scheme to the ECS to display the appropriate CoA (source: Naval Cyber Defence Chair , reproduced with the kind permission of the author).	112
IV.40	Display of the CARMEN plugin in the ECS in the absence of spoofing (source: Naval Cyber Defence Chair , reproduced with the kind permission of the author).	113
IV.41	Display of the CARMEN plugin in the ECS during spoofing, with the suggested CoA (source: Naval Cyber Defence Chair , reproduced with the kind permission of the author).	113
V.1	Research and development perspectives for the Maritime Cyber Situational Awareness concept (source: personal work).	122
C.1	Pre-exploitation modeling using the MITRE Pre-ATT&CK framework on an advanced attack on an ECDIS system (source: personal work).	151
C.2	Post-exploitation modeling using the MITRE ATT&CK framework on an advanced attack on an ECDIS system (source: personal work).	152
D.1	Evolution over time of the AIVDM sentence (source: personal work).	153
D.2	Evolution over time of the GCGGA sentence (source: personal work).	154
D.3	Evolution over time of the GCRMC sentence (source: personal work).	154
D.4	Evolution over time of the GCVTG sentence (source: personal work).	155

D.5	Evolution over time of the GCZDA sentence (source: personal work).	155
D.6	Evolution over time of the GPGGA sentence (source: personal work).	156
D.7	Evolution over time of the GPRMC sentence (source: personal work).	156
D.8	Evolution over time of the IOTXT sentence (source: personal work).	157
D.9	Evolution over time of the NAGGA sentence (source: personal work).	158
D.10	Evolution over time of the NAVHW sentence (source: personal work).	158
D.11	Evolution over time of the NAVTG sentence (source: personal work).	159
D.12	Evolution over time of the NWGLL sentence (source: personal work).	159
D.13	Evolution over time of the NWMWD sentence (source: personal work).	160
D.14	Evolution over time of the NWMWV sentence (source: personal work).	160
D.15	Evolution over time of the NWVHW sentence (source: personal work).	161
D.16	Evolution over time of the PHKS sentence (source: personal work).	161
D.17	Evolution over time of the PHTRO sentence (source: personal work).	162
D.18	Evolution over time of the RAOSD sentence (source: personal work).	163
D.19	Evolution over time of the RARSD sentence (source: personal work).	163
D.20	Evolution over time of the RATTM sentence (source: personal work).	164
D.21	Evolution over time of the SDDBT sentence (source: personal work).	165
D.22	Evolution over time of the SDDPT sentence (source: personal work).	165
D.23	Evolution over time of the SPBDT sentence (source: personal work).	166
D.24	Evolution over time of the SPDPT sentence (source: personal work).	166

List of Tables

III.1	SOC needs <i>vs</i> constraints of the maritime world	36
III.2	Examples of KU and KT data for air context SA elaboration	42
III.3	Examples of KU and KT data for maritime context SA elaboration	44
III.4	Examples of KU and KT data for CSA elaboration	50
III.5	Examples of similarities of KU and KT data for the elaboration of air, sea, cyber SA	50
III.6	Needs in terms of technical data, abstraction and global SA integration for the three different levels of MCSA	55
III.7	Comparisons for detection	59
IV.1	α directing coefficient of the half line representing the <i>stats.capture.kernel_packets</i> parameter on the four industrial loops	86
IV.2	Data timeliness following a satellite outage simulation (source: personal work).	90
IV.3	Characteristics of the NMEA traffic capture files	104
IV.4	Characteristics in period and frequency	106
IV.5	Examples of characteristic sentence sizes	107
IV.6	Talkers and sentence IDs of the <i>cap1</i> , <i>cap2</i> and <i>cap3</i> files	107
IV.7	Examples of cyber talker ID and sentence ID cyber and suggested COA within the CARMEN project	111
A.1	Main differences between Information Technology and Operational Technology. Source: NIST 800.82	142

Acknowledgments

First of all, I would like to thank my PhD thesis supervisors: David Brosset and Jacques Simonin, as well as my PhD thesis director Yvon Kermarrec. Their support and vast experience were particularly valuable for me to take the right course from the start of this thesis and to make me choose the proper research themes. Their experience in research activities was also very important during the reviewing processes of this thesis and of the published articles. Thank for helping me sailing serenely on the fabulous oceans of research!

Thank you also to the reviewers of this manuscript, Nora Boulahia Cuppens and Joaquín García-Alfaro, for accepting this mission despite an already saturated schedule at the end of 2020. I would also like to thank the members of the thesis jury for their presence in the particular sanitary circumstances due to the pandemic, for their careful reading of this work, as well as for the remarks they will address to me and which will allow me to improve myself for my future research.

I also thank the French Navy, my employer, for allowing me to progress throughout my 27 years of service. Difficult to quote everyone, but I especially thank Fabrice, Daniel, Mathieu, Julien, Olivier, Thomas, Eric, Erwan, François-Régis, Xavier, Gladys, Emmanuel, Lionel, Fabien, Loïc, Florian, Alexis, Maxime, Cédric, Laurent, François, Jean-Eudes, Hervé, Marion, Dominique and all of you, who, on a daily basis, work in silence and discretion to improve the cybersecurity of the Navy. Thank you for your support and your trust: this thesis work could have impacted my availability to work and discuss with you on a daily basis, but it was for a good cause! Keep the course!

Thank you to those who believed in the interest of creating a cyber defense chair dedicated to the naval world at the French Naval Academy, as well as a Post-Master's Degree for maritime and port cybersecurity, I am thinking in particular of the Admiral Hello, the Vice-admiral Pagès and the Rear-admiral Baudonnière. I would particularly like to thank

the members of the Cyber Defense Chair of the French Naval Academy, whose good humor and endless advises have strengthened me in my daily work. Thank you also to the other doctoral students: Clet, Pedro, Guillaume, Bastien, Arthur, Étienne, Benjamin, Xavier, Maël, Paul, Nico, Douraid: it has been a pleasure to spend these three years with you, too often part-time. Thank you very much Mallorie and Perrine, our successive communications officers, for your dynamism and for having contributed so much to the influence of the Chair! Mickaël, Maxence, you have done a superb job of turning a stack of boxes into a functional and realistic cyber-range! I would also like to thank the Research Institute of the French Naval Academy, the teachers and researchers of the Academy as well as the staff of the Naval Academy who provided most of the support for this thesis.

Thank you to Mickaël Hauspie and Frédéric Cuppens for their participation in my individual thesis monitoring committee and for their wise suggestions. You both allowed me to clarify my sometimes confused thinking. Thank you also to the staff of IMT Atlantique Bretagne Pays de la Loire and in particular to the UMR CNRS 6285 Lab-STICC laboratory for your welcome during these three years. Invited during the scientific days, you always answer positively and, by your bibliographical and academic remarks and advices, allowed me to advance my research.

Finally, I of course thank my family, who somehow accepted the very important schedule sacrifices linked to this late thesis. To my children: yes your father will be a doctor but no, he will not wear a surgeon's gown at all and will not be able to immunize you against screens (or whatever!). He will try to continue to be the same, hopefully less tired! Thank you also to all those who supported me with their endless question on the date of my defense, helping me to properly set the priorities in my agenda! Thank you in particular to the Djeuns, the Colonel and the Belgian and to Frédéric for their support!

These thanks cannot end without a thought for my first fan: my mother! Her presence and encouragements were for me the founding pillars of the man that I am, of my values and my commitment in life.

In this world where everything is speed, urgency, stress, performance, reorganization, notation, virtual, this thesis work also had a cathartic vocation, bringing me an undeniable and beneficial perspective and a projection on fields in full construction. It also allowed me to reflect on myself, my mid-life journey and my desires for the future.

CHAPTER I Introduction

We are drowning in information, while starving for wisdom. E.O. Wilson, 1929

February 20th, 2020: an ocean liner with 8,000 passengers and crew transits the Mediterranean Sea. While passengers are busy with their leisure activities, the ship is subject to a "digital hostage-taking". The shipowner's crisis cell is activated, the highest authorities of the State are notified: the captain can no longer maneuver and all the screens of the bridge and the machine's control room display an unequivocal message: the owner must pay a high ransom or the ship will run aground on the coast at full speed. Is it the screenplay from a bad Hollywood movie? Perhaps. An already occurred incident? We will probably never know. A feared event present in the risk analyses of these ships? We hope so.

In France, maritime transportation is one of the sectors of activity of vital importance (SAIV)¹. Over the past few years, the improvement of the cybersecurity level in sectors of vital importance appears noticeable, following the regulatory work of the *Secrétariat Général de la Défense et de la Sécurité nationale (SGDSN)* and more particularly the dynamism of the French National Security Agency (*Agence Nationale de la Sécurité des Systèmes d'Information (ANSSI)*). From now on, processes exist to ensure the reporting of a maritime cyber event within the framework of the *Coopération Navale Volontaire (CNV)*². However, several studies agree to underline that the level of cybersecurity of the maritime sector remains insufficient [DiR15, Tam19]. This observation, shared by French, European and international cybersecurity players, challenged us.

Due to the lower cost of transport it offers, the maritime sector is the world's leading

¹http://circulaire.legifrance.gouv.fr/pdf/2014/01/cir_37828.pdf

²Read in particular <http://www.sgdsn.gouv.fr/uploads/2018/08/20180627-instruction-interministerielle-nxx-230-surete-maritime-portuaire.pdf> and <http://www.sgdsn.gouv.fr/uploads/2019/05/20190429-sgdsn-iim-165-cooperation-navale-volunteer-vf.pdf>

mode of transport for consumer goods: thus, in 2019, it accounted for more than 4/5th of freight traffic worldwide in volume [SHJ⁺19]. Cyber vulnerabilities exploited by attackers could compromise entire sectors of the economy by causing denial of access to many raw or manufactured materials. This observation has strongly motivated all of our work: we have therefore focused our research efforts on the design of models and solutions enabling to practically improve the detection and analysis of cyber events in the maritime environment and strengthen maritime security as a whole.

Context of the study

With the digital technology revolution, the maritime sector has been experiencing, for the past ten years, a deep and long-term transformation of ships, port infrastructures, equipment and operating processes. The sector is changing like other sectors before it, but at its own slower pace, linked to the long life cycle of parts of ships and port infrastructures. If the bridge of older ships still retain a mix of physical switches, buttons, analog screens and computers, recent port infrastructures and new ship constructions are already much more digitized and the first sea trials of autonomous ships have become a reality [Lev17]. This major development is also found in the military naval sector. In a modern navy, no large-scale operation could now be carried out without the use of digital technology. As a strategic advantage for intelligence and anticipation, deterrence, protection, prevention and intervention, digital technology has tremendously increased the capacities of our armed forces.

The maritime information systems enabling these digitized processes present a large attack surface, in particular due of their interconnection, and are equally well exposed to attacks originating from sources of threats linked to cybercrime or to state-level actors. Sometimes poorly designed and/or exploited, weakened by the vulnerable nature of software solutions, maritime systems can suffer from organizational, environmental, human, technological or technical weaknesses. When these vulnerabilities are combined and exploited by malicious people or by accident, the functioning of information systems and facilities can be directly affected. The spectrum of consequences is vast, from a simple mild viral infection to cyber physical, environmental, human impacts, to the destruction of facilities, or even to the compromise of the conduct of military actions.

Cybersecurity tools are gradually being deployed within certain maritime information

systems. They are based on three pillars: cyber protection, cyber defense and cyber resilience. Cyber protection corresponds to organizational and technical in-depth defense measures which make it possible to reinforce a system during the design and operation phases, to guarantee a high level of security. Cyber defense brings together all the means that make it possible to ensure the detection and response to a cybersecurity event³. Cyber resilience provides the means of patch management processes, to provide regular updating capabilities, fixing system vulnerabilities and, *in fine*, continuity and recovery of activity in the event of a cybersecurity event. If the contribution of this thesis focuses primarily on cyber defense, part of the research findings and information obtained can contribute to improve the cyber protection and cyber resilience of maritime systems. They can also be extrapolated to serve other fields of employment, such as autonomous vehicles or connected objects. The significant improvements in the collect, analysis and understanding of cyber events provided by our model also makes it possible to respond to other fields of anomaly detection, such as the premature wear and tear of cyber physical installations, or to uncover the value of operational data, locally or remotely.

Theoretical framework of the research

The difficulties to implement traditional cyber security measures on board ships and maritime infrastructures mostly find their origins in three issues specific to this environment. The first is linked to the very long life cycle of ships. A ship or an industrial system present on a port infrastructure is designed to be operated for several decades, with a possible overhaul at mid-life. This long time scale clashes with the daily operational reality to optimize the operation of the ship, but also with its design phase which cannot be delayed, under penalty of significant economic consequences. The multiplicity of players (shipyards, operators, maintenance staff, etc.) also complicates the optimization of this life cycle. In a highly dynamic cyberspace, with a daily evolution of threats and vulnerabilities, the obsolescence of systems is therefore a reality with which we must learn to live. It is thus often impossible to update and secure all the systems to face threats: it is necessary to identify other solutions to protect them. The second problem is linked to the integration of heterogeneous maritime systems and their necessary interconnection for operational use: the ship is a system of complex systems. Acquiring a global and exhaustive vision of its systems from the design stage

³By cybersecurity event, we mean an unexpected and non-compliant situation in cyberspace, voluntary or involuntary, which could have an impact on the operation or integrity of the system, the data it contains or the installations it operates.

throughout the ship life cycle is a wise process[For13]. Classic cyber protection solutions are also difficult to adapt to the particularly constrained environment of maritime information systems, because their design does not take into account their peculiarities. Finally, except in very few cases, no specialist in IT, network and even less in cyber defense is embarked on board a ship. The maritime sector is, therefore, faced with a multitude of obstacles, organizational and technological, endogenous and exogenous, hampering the integration of effective cyber security measures on board increasingly complex and digitized ships.

No solution can be considered as perfect, due to lasting, structural and sometimes exogenous difficulties: it is therefore unreasonable to hypothesize that, one day, the ship will be less digitized or that cyber experts will be present on board all ships on a permanent basis. As part of this thesis, while deliberately following a holistic approach, we wanted to focus on the detection, collection, analysis and understanding of cyber events in a maritime context. This capacity could indeed make it possible to anticipate, detect quickly and provide an understanding and a response to the sequence of events to avoid the potential disaster scenario presented at the beginning of the introduction. Indeed, while work has already been carried out or is continuing on subjects related to cyber protection or resilience of information systems, the centralization of cyber defense events in a maritime context for the purposes of detection remains an area of concern where research is still underdeveloped for the moment. This thesis, which focuses on the challenges of maritime cyber defense, provides a proven response in the field which has been the subject of several publications at international conferences. By implementing cyber monitoring of maritime information systems, real-time threat detection and risk assessment take a practical dimension, facilitating the defense and processing of cyber events and improving overall maritime security.

Research questions

In the maritime context, the research problem that motivates this thesis is above all to improve the quality and speed of detection, understanding and reaction to a cyber threat. To answer this problem, three research questions were identified to support this thesis.

1. RQ1: Collection architecture and modeling: what are the processes to be designed, developed and implemented to enable anomaly detection on remote and heterogeneous systems which have constrained telecommunications circuits? What global modeling could be used to create knowledge from the data collected? The first objective pursued

in this thesis will be to demonstrate that the concept of situational awareness applied to maritime cybersecurity makes it possible to respond to these challenges. Our work will also aim to attest that the obstacles linked to the collection, standardization, fusion and visualization of multi-sensor data can be solved.

2. RQ2: Analysis and representation of data: what are the achievable methods of processing, analysis and presentation of data to have a relevant and understandable visualization of the situation of a fleet of vessels for the purposes of monitoring, projection, sharing and decision support? Faced with the high volume of data collected by multiple sensors, our work must identify the relevant visualization and analysis tools in order to have an overall understanding of the context to facilitate human decision-making.
3. RQ3: Detection of anomalies in a maritime context: the maritime world uses protocols that are sometimes not secure by design, not scalable and vulnerable to a multitude of attacks. How can anomalies be detected in this context? We propose to apply detection technologies to a maritime-specific systems and standards and to present detection information in a simplified manner to their operators.

Adopted approach and contributions of the thesis

As part of the first research question on collection architecture and modeling, we focused our work on the notion of situation: a man who faces events, circumstances and relations between objects finds himself facing a situation. Research related to the perception, understanding and projection of the situation made it possible to retain the concept of Situational Awareness (SA), already used in other fields of research and application. We used this work to define the concept of Maritime Cyber Situational Awareness (MCSA). First of all, we have precisely modeled the different processes that lead to the state of knowledge that allows the MCSA. Subsequently, based of the information needed for the development of these processes, we have defined an architecture tailored for the maritime world in order to ensure a proper on-board collection and data transfer ashore. In order to evaluate our work, we directly participated in the creation of an experimental platform, one of the objectives of which was to integrate the specific characteristics of the environment, such as the constraints linked to satellite connections or the variety of embedded systems. Finally, to allow adequate decision-making, this information must show a high level of quality, based on several criteria that we have also precisely defined. The concept of MCSA for the maritime sector and

ships of the future has been the subject of several articles and publications in international journals and conferences.

To answer the second research question on data analysis and representation, we worked on the anomaly detection techniques in a maritime context, the understanding of the attack path and the contribution to remediation, by helping the user visualizing and understanding the events in their context. Our research work has in particular provided answers on the necessary evolution on network sensors to allow the realisation of dynamic graphs. We have also developed dynamic and contextual dashboards, making it possible to display tactical situations according to several dimensions, meeting the needs of the operators and the expected level of abstraction. The mapping and identification of dynamic attack paths in a context of real-time massive data processing represent important contributions for the field of research. This work has been the subject of several scientific publications and improvements to the source code of sensors.

Finally, the third research question related to anomaly detection in the maritime context allowed us to define a real-time detection architecture on specific systems and protocols. We were able to identify and demonstrate the weaknesses of certain standards used on board ships. We designed, experimented or directly contributed to an advanced system offering operator-level detection and alerting capabilities.

The commitment of the various actors solicited within the framework of this thesis and a favorable context allowed us to validate all of these contributions by carrying out an operational implementation of our work on a dedicated and realistic architecture: the Naval Cyber Range of the French Naval Academy, developed in parallel with this thesis. The particularly valuable experience feedback was essential for us to validate our methodology, our proposals and verify the conformity and reality of our contributions. In fact, in parallel with the creation of the Naval Cyber Range, we have implemented a global cyber monitoring architecture similar to what could be deployed on board a ship and in a shore center to ensure the cybersecurity of a ship or a fleet of ships. The instrumentation of the Naval Cyber Range as a prototype allowed us, on the one hand, to generate cybersecurity metadata and collect them in order to carry out our quality measurements. On the other hand, it enabled to carry out our work on analysis, visualization and anomaly detection on maritime systems in order to verify the conformity of our research work in a very realistic manner.

The research fields treated in the framework of our work are summarized in figure I.1.

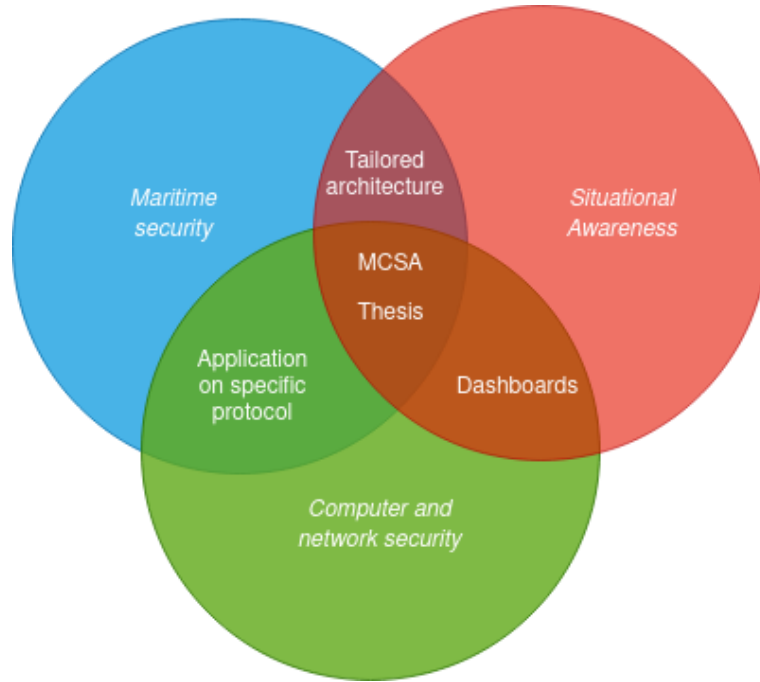


Figure I.1: Research fields of the present thesis (source: personal work).

Plan of the manuscript

The manuscript is composed of five chapters (figure I.2).

After having drawn up a brief state of the art of the maritime world, its information systems and their cybersecurity in Chapter 2, we will describe in Chapter 3 the hard points encountered in the design of a maritime cyber monitoring architecture, as well as that the contributions of our work to model an architecture which meets the various constraints. In Chapter 4, we will detail the results obtained on an experimental platform and identify the means necessary for the analysis and contextual presentation of data as a decision aid, as well as the contributions identified in the field of data visualization for a maritime cyber security center. Finally, we will present our work and results relating to the anomaly detection in the maritime context. In Chapter 5, the conclusion will sum up the contribution of our research work on the identified scientific obstacles and will open to several research perspectives, in particular on ships and more generally, on autonomous vehicles and on the applicability of our work to other constrained fields, such as the Internet of Things.

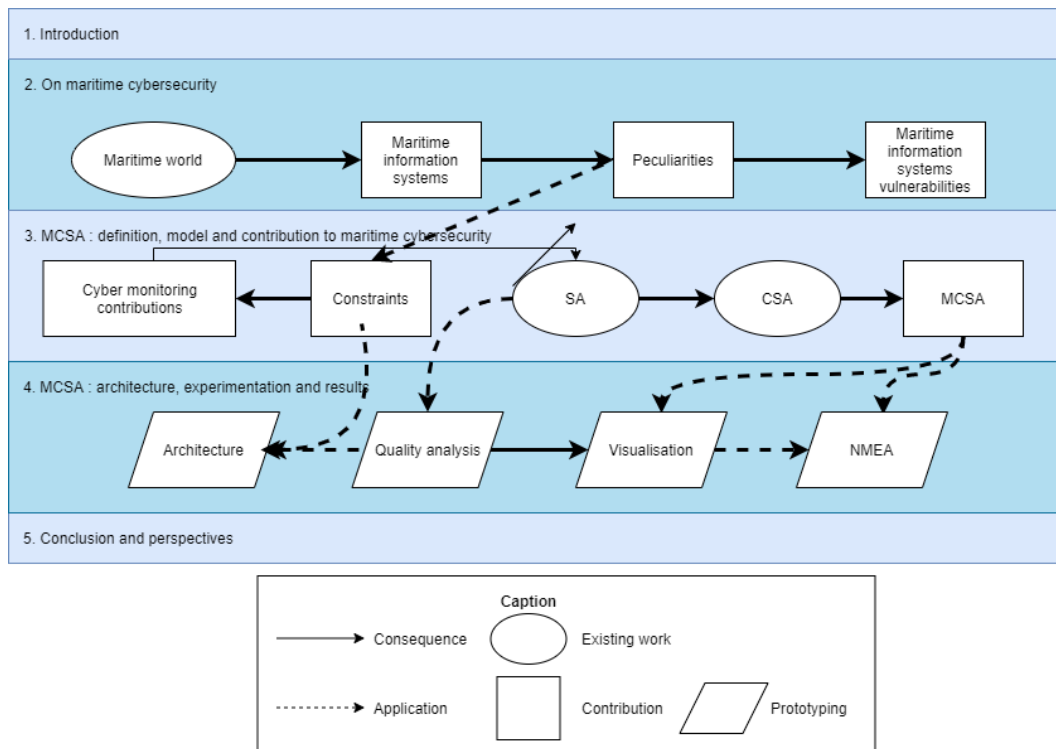


Figure I.2: Description of the relations between the chapters of the manuscript of the present thesis (source: own work).

CHAPTER **II** --- On maritime cybersecurity

In this chapter, we recall the main characteristics of the civil and military maritime sector and the reasons for its strategic importance for France. The digital transformation of the sector is then detailed, specifying the origins of this transformation and explaining the use of maritime information systems. Then, the peculiarities of maritime information systems are identified and some prospects for short and medium term developments drawn up. Finally, we list the vulnerabilities and risks of the domain.

II.1 The strategic importance of the maritime sector in France

After having painted a portrait of the global maritime sector in its geographical and geopolitical aspects, we will approach the specific aspects of the maritime sector in France, under both civilian and military aspects.

II.1.1 The maritime world

Oceans cover 71% of the earth's surface. Historically, humans have exploited them for the transport of goods and passengers, but also for fishing activities and oil and gas production. The use of maritime capacities for the transport of large quantities of goods constitutes an efficient and economical means. This is one of the reasons why 90% of the volume of freight traffic uses the sea route [oS]. To further reduce the cost of this mode of transport, the maritime sector designs and then operates ever larger vessels. For example, the largest container ships built today are capable of carrying over 23,000 containers [Ins]. The figure II.1 shows the evolution of the tonnage of ships over the last thirty years. The increase

in overall tonnage over recent years is particularly significant. Thus, a recent study by the Organisation for Economic Co-operation and Development (OECD) indicates remarkable increases in tonnages over the period 1996 to 2015, in particular for container ships (+80%), bulk carriers (+60%) and passenger ships (+30%) [Mer15].

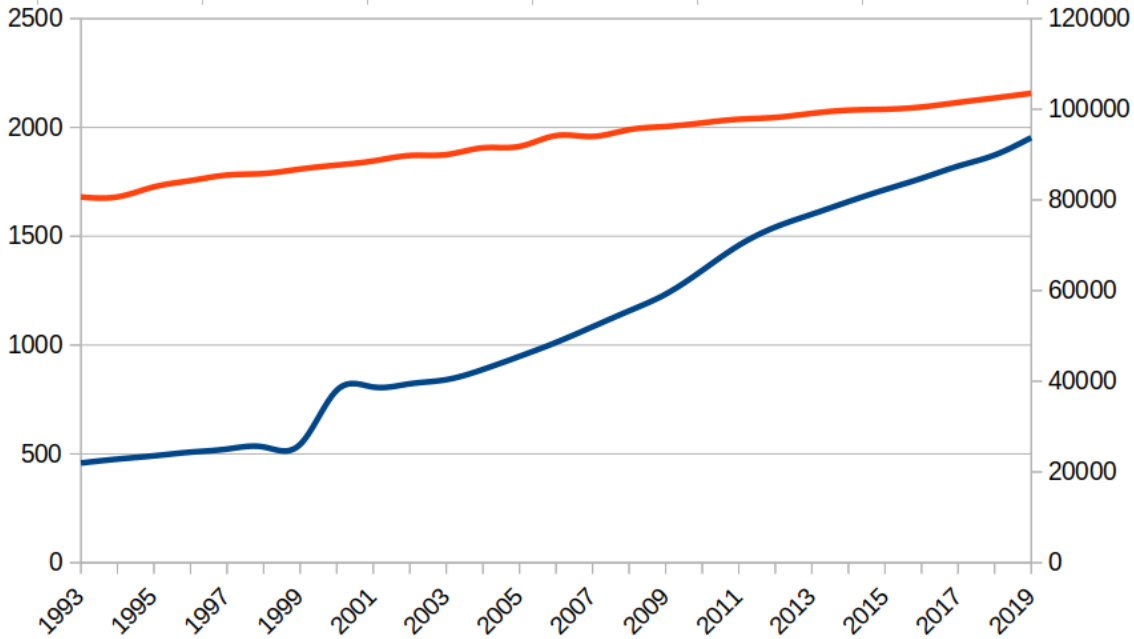


Figure II.1: Evolution of the number (left) and of the gross tonnage (right) of ships on the 1993-2019 period (source : *Lloyd's Register Foundation*, © *Lloyd's Register Foundation* 2018. All rights reserved).

While transatlantic passenger transport declined with the boom in the aviation sector in the 1960s, maritime passenger transport gained a new lease of life with the development of cruises. Here too, the race for gigantism is essential, with ships that can accommodate up to 9,000 people on board, passengers and crew members [Cen]. The figure II.2 offers a comparison of dimensions between an Airbus A380 and a few ships of remarkable sizes: the nuclear aircraft carrier "Charles de Gaulle", the transatlantic ocean liner "Queen Mary 2" and the container ship "Antoine de Saint Exupéry" of CMA/CGM¹.

Fisheries, oil and gas resources are essential to our daily life. 90% of the quantity of crude oil, essential to the needs of France, transits by sea and France also depends on the sea for the transport of many strategic raw materials. In the years and decades to come, the gradual depletion of certain resources and competition in access to raw materials will lead to ever greater geopolitical tensions, including at Europe's maritime borders. Maritime

¹Respectively commissioned in 2007, 2001, 2004 and 2018

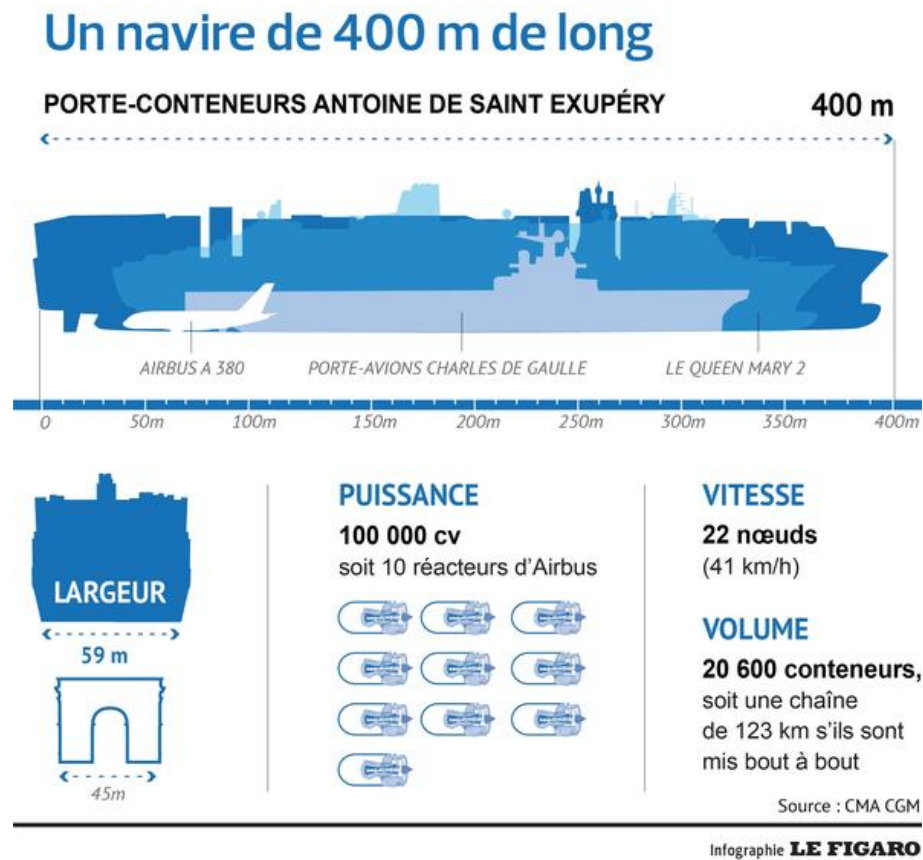


Figure II.2: Relative dimensions of modern ships (source: CMA/CGM, infography: [Le Figaro](#), © Service Infographie / lefigaro.fr / 09/06/2018, reproduced with their kind authorization).

transport could therefore represent a prime target and acts of piracy increase, whether political, economic or terrorist.

Finally, the maritime sector contributes directly to the daily use of digital technology: 95% of international Internet and phone communications pass through submarine cables. With the explosion in the number of terminals, especially mobile ones, and the development of the Internet of Things and cloud computing, the volume of data passing through the Internet is constantly increasing, causing countries to really depend on the submarine cables [Bla18]. The loss of essential services and the impossibility of exchanging data following the break of one of these cables would therefore have major impacts on the economy and the functioning of a country and would generate significant geopolitical tension [Cla16].

The maritime sector has virtually always known acts of piracy. Novels and history books are teeming with tales of maritime piracy, sometimes even mandated by the state.

If we talk less about it today, piracy remains a daily reality for many sailors. Maritime routes are well known, as are the existence of vital straits and canals. The figure II.3 thus shows that the Panama or Suez Canal, the Straits of Hormuz, Bab el Mandab and Malacca are essential access routes for the transport of minerals for many countries including France. Maintaining free access to sea lanes and straits therefore represents a major stake for the economies which depend on them [dS16]. Beyond political and diplomatic actions, this free access is guaranteed by the use of a deterrent military maritime presence. After a very relative period of uncertainty, we can now see that many emerging economies are significantly strengthening their maritime military capabilities, mainly in an arc stretching from the Eastern Mediterranean to Australia, via India and Asia.

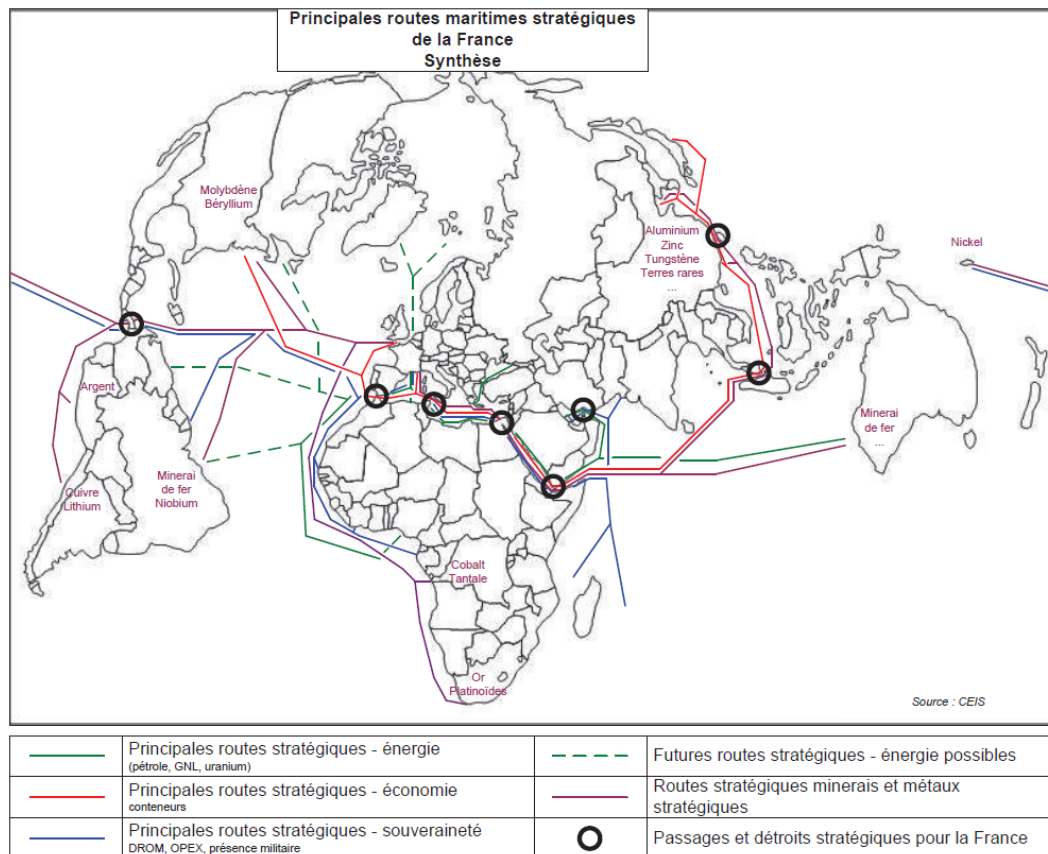


Figure II.3: Strategic maritime routes for France (source: [Étude "Impacts sur l'économie française de la fermeture d'un ou plusieurs détroits maritimes majeurs"](#) of the Compagnie européenne d'intelligence stratégique, Janvier 2016. © CEIS, All rights reserved, reproduced with their kind authorization).

The global maritime sector is therefore changing to face the new challenges of freight and passenger transport. The dimensions of ships are increasing sharply to meet the demands

of the globalized economy. The underlying issues of free movement and sovereignty over the oceans remain essential for the political balance and the daily life of billions of women and men (food, manufactured goods, etc.).

II.1.2 The French civil maritime sector

France is known to be a important maritime power [Lab20] [Bri19]. The relatively low knowledge of maritime issues by most of the French can be explained in particular by the modest size of the continental hexagon, by the magnetism of the capital to the detriment of the ports and by the relative ignorance of supply routes in a globalized economy. However, taking into account the overseas territories, France has 7,000 kilometers of coastline and has the second largest maritime domain in the world, after the United States and before Australia, with an Economic Exclusive Zone (EEZ) of 11 million square kilometers. The French maritime sector, reputed to influence 14% of the Gross Domestic Product (GDP), comprises 340,000 employees and generated 83 billion euros in production value for the country in 2018-2019². The maritime sector also plays a strategic role in allowing France to access many minerals such as, for example, nickel, silver, copper, lithium, beryllium, aluminum, zinc, tungsten, iron and cobalt (figure II.3).

The French maritime world is rich as a host of complementary activity sectors [Vei12]. Let us first of all evoke the major maritime commercial ports, the two best known being Marseille-Fos and Le Havre-Rouen-Paris (HAROPA). These two strategic locations are sometimes called *smartports*, due to the major telecommunication, automation and, more generally, digitization projects being carried out there to improve logistics flows and security. Other major seaports of importance such as Dunkirk, Nantes-Saint Nazaire, La Rochelle and Bordeaux deserve to be mentioned. Some have specialized in shipbuilding and repair, such as Saint-Nazaire (Chantiers de l'Atlantique) and Brest. It should also be noted the establishment of several petrochemical sites and the presence of major *hinterlands*³, as in Marseille-Fos, Havre-Rouen-Paris or in the Rhône river valley (Port Édouard Herriot, near Lyon). French shipowners are particularly represented internationally, whether in the transport of goods, with shipowners such as CMA CGM or Dreyfus, or through cabotage and river traffic which are revitalizing, particularly on the Seine, the Rhine and Rhône. Most of the passenger transport volume is carried out on cross-Channel traffic (Brittany Fer-

²Read https://www.cluster-maritime.fr/economie_maritime/

³The *hinterland* corresponds to the mainland attraction and service area of a port.

ries) and between Corsica and the mainland. The cruise sector, which was booming before the COVID-19 pandemic, is internationally recognized thanks to operators such as Ponant for exceptional cruises. In a competitive sector, fishing is an important maritime activity in France, in particular for the ports of Boulogne-sur-Mer, Le Guilvinec, Erquy, Port-en-Bessin, Concarneau, Roscoff, Sète and Lorient. Scientific and oceanographic research and exploration are active sectors, with reputable organizations such as Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER) or Compagnie Maritime d'Expertise (COMEX), not to mention the environmental mapping missions carried out by the Service Hydrographique et Océanographique de la Marine (SHOM). Oil and gas exploration and transport are represented by several companies such as Bourbon, Total Energies or Elengy, as well as the design, installation and repair of submarine cables with Orange Marine or Nexans. The yachting sector is renowned for the quality of shipbuilding and the number of boats in marinas in French ports, but also for its ambition and its recognized results in ocean racing. Finally, more recently, France has positioned itself in the Marine Renewable Energies (MRE), with the development of offshore wind farms, tidal turbines, as well as construction, maintenance and onshore operating capacities, with companies like Sabella and the France Énergies Marines institute.

Even if competition is strong in Europe for the transport of goods, with the ports of Rotterdam and Antwerp, and despite the impacts linked to the pandemic, the French maritime sector can be considered as rich, very varied and endowed with a real potential.

II.1.3 The French military maritime sector

The French Navy's main objective is to carry out four defense missions. The first is deterrence: it is a question of ensuring, on the one hand, the permanence of oceanic nuclear deterrence⁴. And, on the other hand, airborne deterrence tactical capability with the Rafale marine aircraft on board the Charles de Gaulle nuclear aircraft carrier. For more than 40 years, with the construction of the first ballistic missile submarine (*Sous-marin Nucléaire Lanceur d'Engins (SNLE)*) and the first operational patrols, the permanence of strategic oceanic deterrence has never failed. Second mission: the protection of populations, territories and the EEZ, daily objective of many actors of the French Navy, on the seas, in the air and in overseas territories. The attraction of the halieutic resources of the second EEZ in the world involves its share of looters, off Guyana as well as in the South Atlantic, against which it is

⁴Deterrence can also take non-nuclear forms.

necessary to fight permanently. The third mission is to prevent maritime risks, for example by ensuring regular patrols and the pre-positioning of naval forces in conflict zones. Finally, intervention: when necessary, the French Navy has forces capable of intervening militarily, far and long, whether for the recovery of nationals or for military strikes, a capacity that has been proven on several occasions during recent operations.

The second function of the French Navy covers the action of the State at sea. The Maritime Prefects, general officers of the French Navy, play an essential role in this function. Beyond a role comparable to that of coast guard, the Maritime Prefect also represents the State and can give notice to a shipowner whose ship would present an immediate danger to the French coast, for example in the event of major damage of propulsion⁵.

The French Navy today comprises more than 39,000 sailors, military and civilians, 72 offshore vessels, 200 aircraft, 10 nuclear submarines, 2,000 marines and commandos and 1,000 maritime gendarmes working on 30 smaller vessels [nat]. In mainland France, most of the French Navy's ships are built, maintained and moored in the four major military ports (Brest, Cherbourg, Lorient and Toulon). The SNLE are moored and maintained on the operational base of Ile Longue, near Brest. Finally, the Navy is widely present in overseas territories as well as abroad, through partnerships signed with allied countries.

Faced with strategic maritime challenges, France therefore has a top-ranking Navy, in full renewal and whose operational excellence has been regularly recognized, including during recent military engagements, particularly in the eastern Mediterranean.

II.2 The digital transformation of the maritime world

The maritime environment, like other sectors, is experiencing an unprecedented wave of information technology use. The construction and operation of modern ships makes extensive use of "new" technologies. Indeed, conning, maneuvering and operating ever larger and more efficient ships with optimized crews⁶ requires the use of digital technologies [Tam18b]. Likewise, when the ship is docked in a port, the efficiency of logistics processes relies on computer and telecommunications technologies in order to unload goods quickly with complete safety and to ensure efficient maintenance operations. The construction and renovation of

⁵which cause could also be cyber

⁶for instance, a recent multi-mission frigate in France has 108 crew members, compared to double for a previous generation anti-submarine frigate.

ships, as well as land-based infrastructures, are gradually but definitively switching to digital, without any global plan having truly coordinated or formalized this transformation or measured its consequences: in some ways, like in other sectors, the maritime world is largely subjected to what the law of the market dictates.

The - relative - slow pace of this digital transformation can be explained, in particular for ships, by a long life cycle. It is not uncommon for a ship, after a design phase of 15 years⁷, to have a lifespan of over 40 years. Ships built in the 1980s can therefore still be found on the seas. These ships, although gradually switching digital, still rely heavily on analog. It is especially from 2050 on that the maritime sector will probably have completed its first major digital migration, with the disappearance of old ships and the omnipresence of highly digitized ships and autonomous ships.

Digital technology constitutes an indisputable vector of progress for the maritime environment, for the safety of people, the monitoring and safety of cargoes and passengers and therefore for global economic development. This transformation is taking place in a particular regulatory context: if international maritime regulations only just mention lightly cyber issues, the standards of the maritime sector often impose specific requirements in terms of resilience, in particular for cyber physical installations, that is to say - industrial installations controlled by data processing to carry out physical operations of production or maneuvering by means of automation, sensors and actuators. These installations are therefore generally redundant, but not necessarily secure from a cyber point of view.

Regarding the development of autonomous ships (also called Maritime Autonomous Surface Ship (MASS)), the International Maritime Organisation (IMO) has recently set strict regulations concerning their sea trials and defined four degrees of autonomy. The first degree corresponds to a ship that uses automated processes with an autonomous decision-making capacity: sailors on board operate and control the systems and functions of the ship. Some functions can be automated and sometimes unsupervised, but the sailors on board stand ready to take over when needed. The second degree extends to ships that can be remotely controlled, but which still have sailors on board: the ship is controlled and operated from a remote position, but the sailors on board can take control of the ship and operate the on-board systems when needed. The third degree does not foresee any human presence on board: the ship is totally controlled remotely. Finally, the fourth degree concerns fully autonomous ships where the ship's operating system makes decisions and determines

⁷For more complex ships, such as combat vessels.

its actions in complete autonomy⁸. The transition to semi-autonomy or full-autonomy of ships will require a strong use of algorithms and digital technology in the years to come, as well as ambitious regulations on the part of the IMO on the related issues such as resilience, autonomous vessel decision-making and cybersecurity.

The sustainable digital transformation of the maritime sector is therefore a reality: digital is above all a vector of strength, quality, profitability and safety. Resilience remains at the heart of the concerns of seafarers, who are fully aware of the risks associated with the sea and the importance of saving human life. However, the lack of integration of cybersecurity on board and this new dependence of ships on digital and, sometimes, on land, could lead to major consequences which could go as far as "decision paralysis" [Goe15].

II.3 The "marétique"

The Blue Book of the Marétique cluster, published in 2012, gave the first definition (II.1) of the French word "marétique"[mf], without English equivalent to our knowledge.

Definition II.1. *"Marétique" brings together all the computer and electronic systems used in the management and use of operations relating to maritime, river and port activities.*

We propose to use the term "Maritime Information Systems (MIS)" in this thesis to evoke the sole computer component of "marétique". As we saw in the previous chapter, the MIS make it possible to multiply the capacities of the maritime world and to improve its security. In addition to a spatial separation of "marétique" ("marétique" on board ships, MRE and *offshore*) and that present in land maritime infrastructures, the National Institute of Standards and Technology (NIST) suggests a classification in Information Technology (IT) and Operational Technology (OT) [Sto11], recalled in appendix A, which divides information systems in two, based on technical characteristics, generally discriminatory. However, these systems converge technologically and this difference sometimes tends to disappear [Gar18].

II.3.1 Onboard maritime information systems

The MIS are numerous and cover practically all the essential functions of a ship. For most ships, civil and military, the functions that define the MIS are common.

⁸<https://www.imo.org/en/MediaCentre/HotTopics/Pages/Autonomous-shipping.aspx>

The IT firstly includes communications to the outside world: this category includes the extension of services and networks from shore to land (telephony, connection to the Internet or to the company network, videoconferencing...), using Global System for Mobile communications (GSM) / 3G / 4G / 5G networks near the coasts or satellite or radio links in the high seas. Internal communications include general broadcasting on board the ship, alarm broadcasting systems, telephony and intercom systems, Internet Protocol (IP), Wireless Fidelity (WiFi) or Digital Enhanced Cordless Telephone (DECT), equipment such as routers, switches, Virtual Private Network (VPN), Virtual Local Area Network (VLAN), various security solutions (gateways, sensors, firewalls) or logging tools. This category also includes entertainment systems for crew and passengers, with Internet access at sea via a WiFi or wired connection, any local GSM relay, dematerialized medical monitoring, payment systems, passenger boarding control, hotel management (Property Management System (PMS)), as well as television on demand. IT also covers corporate intranets, messaging and local servers. Finally, access control systems, Closed Circuit TeleVision (CCTV), bridge security (Bridge Navigational Watch Alarm System (BNWAS)), alarm systems towards the earth, such as the Shipboard Security Alarm System (SSAS) and the paging and geolocation systems can also be grouped under the term IT.

OT concerns the ship's control systems (*cf* figure II.4): electronic nautical charts systems (Electronic Chart Display Information System (ECDIS), RADARs, sounders, geographical positioning by satellite (Global Navigation Satellite System (GNSS), such as Global Positioning System (GPS), Galileo, GLONASS), time reference servers, dynamic positioning systems (Dynamic Positioning (DP)), meteorological and nautical situation awareness tools (Automatic Identification System (AIS), Global Maritime Distress and Safety System (GMDSS)), maritime flight-recorder-like solutions (Voyage Data Recorder (VDR)) and other logging and alerting tools. OT also integrates the operation of the platform: electricity or chilled water production systems, propulsion management, steering management (helm, stabilization), opening and closing of watertight doors, diesel fuel supply, treatment of gray and black water and air, refrigerators, alarms and fire safety or flooding. The cargo management systems are also concerned: Cargo Control Room (CCR), as well as information systems for monitoring the goods and the loading plan and the ballast and purge management systems.

Warships have special OT systems, such as combat systems with sensors and effectors such as sonar, radar, gunnery and missile control systems, electronic warfare, nuclear reactor control systems (for ships fitted with such installation), but also with specific military navigation systems such as, for example, inertial navigation units.



Figure II.4: The bridge of the research vessel *Sikuliaq* (source: [Wikimedia Commons](#), unmodified picture, author: NorthBySouthBaranof, license: [CC BY-SA 4.0](#)).

II.3.2 Ports and offshore installations information systems

Ashore, within ports, but also for shipowners and vessel traffic monitoring organisations, there are also many IT systems, such as those allowing the management of the shipowner's fleet, reservation systems services (purchase of trips, container transport services, Internet and intranet/extranet sites), maritime customs exchange systems or even maritime traffic coordination and maritime information exchange systems. As for OT, we can think of management systems and the loading and unloading systems for fluids, bulk and material (containers), port platform planning and management systems, particularly complex oil and gas platform operating systems and, finally, MRE production equipment control systems.

II.3.3 The peculiarities of MISs

A ship that accommodates 8,000 people on board is a true small floating city which includes, as we have seen, a significant number of information systems. A nuclear aircraft carrier at sea thus combines the peculiarities of a nuclear power station, a military airport, a pyrotechnics system, a hospital, a hotel and catering capacity for 2,000 sailors, a aeronautical maintenance center etc., all deployed hundreds or even thousands of kilometers from the nearest coasts, sometimes in a hostile context and with limited and distant support resources.

MISs therefore also inherits a combination of the particularities of ships, their architecture, their design, in particular use-cases, all in an often harsh environment.

In a cyber context, the peculiarities of MISs that we have been able to identify, essentially for on-board systems, are first of all due to the multiplicity of systems. Whether in terms of quantity, type, manufacturer or technologies, the relatively low level of integration of these systems leads to heterogeneity with significant consequences. The constraints weighing on satellite links are high: the bandwidth is limited and the cost of the link generally depends on the bandwidth or the volume of data allocated. In addition, the permanence of the link is not guaranteed, which can complicate remote maintenance and monitoring, not to mention the impacts related to latency. Then, even in the event of damage to one or more of its systems, the vessel at sea must remain resilient. Maintenance interventions are generally only possible during the ship's port of call in a port. In most cases, no IT expert or technician is present on board the vessel, especially on civilian vessels: the capacities for updating, installing patches and restoring the system are therefore limited. Finally, on land, no integration platform representative of the actual situation on board the ship has generally been designed. In fact, most of the systems have been produced by different companies, and have sometimes been integrated on board by another company. In addition, the evolution and the update of an installation can require a complete new certification before any deployment: the patch management of these systems becomes, consequently, complicated. If, each taken separately, these characteristics may not seem specific, their addition, summarized in figure II.5, contributes to make a MIS and its cyber security so specific. As a result, most off-the-shelf solutions may prove unsuitable for direct on-board application [Jac18].

The role played by the designer of the ship is essential. As an integrator of potentially heterogeneous systems developed by several subcontractors, he has to ensure the ship's overall cyber security and its consistency to face the threat. Even in cases where the integration level is low, the coupling level between systems can be higher than expected. Above all, this means that the sharing of services and common functions - especially cyber security - has not been thought out or optimized: each onboard IT or OT system, being manufactured by various manufacturers and poorly integrated, has functions that are independent in terms of security (network, application, system) from those of a neighboring system, without any homogeneity having been sought. At the same time, all systems on board recent ships are digitized in order to be monitored and controlled remotely. For instance, a food refrigeration installation, previously purely mechanical and analog, is today frequently controlled by a Programmable Logical Controller (PLC) with sensors and actuators, with one or more

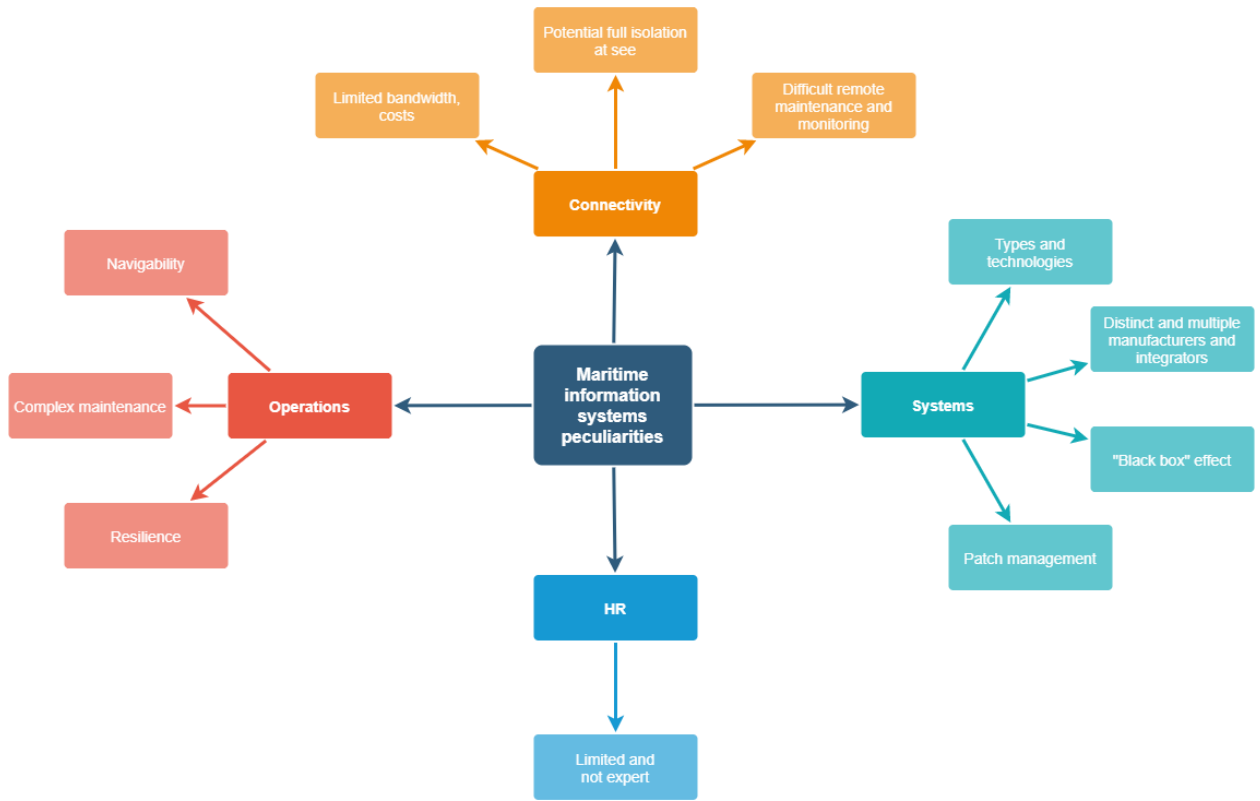


Figure II.5: Peculiarities of MISs (source: personal work).

control-command links to the ship's operating system, or even to shore. The digitization of the ship therefore leads to a dependency that did not exist previously. At the scale of the ship, despite a rather incomplete integration in terms of cyber security, digitization is gradually bringing a stronger functional integration of the systems. By increasing the functional inter-dependencies, it also increases the complexity of the whole and the associated risks, particularly with limited human capacities.

II.3.4 The future of MISs

MISs will continue to support transformations in the maritime sector in the future. The major changes to be expected in the short and medium term mainly concern the development of semi-autonomous and autonomous vessels, the first tests of which took place during the year 2018 and are continuing today, in France, in Europe, and in the whole world. In the decades to come, their development could, in certain cases, limit or even eliminate the use of humans in all or part of certain areas of the maritime sector, substantially increasing the digital attack surface of these ships [Tam18a, Kav18, Kat17, Bol19]. The generalization of

autonomous vehicles (air and sea, surface or underwater) in the maritime sector also applies to the military world, in order to limit human exposure to risks and ease the deployment of sensors, in particular for intelligence purposes. Finally, in the shorter term, the implementation of advanced digital technologies such as machine learning and augmented reality will support people in the maneuvers, operation and remote and local maintenance of ships and infrastructures.

II.4 Cyber security of Maritime Information Systems

After having declined the definitions of cyber security to the maritime world, this section details the cyber vulnerabilities and risks of the sector.

II.4.1 Cyber security and maritime environment: definitions

The term "information system" is subject to many definitions and interpretations depending on people and organizations. In this manuscript, we will therefore use the definition II.2 of an information system:

Definition II.2. *Manual or automated system, such as an automatic data processing system, computer system or computer network, relying on technical infrastructure and composed of people, machines and methods and which is organized to perform collect, processing, transmission and dissemination functions of data which represent information.*

Figure II.6 specifies the different views that can be produced of an information system, from a functional, application and technical point of view.

In the maritime context, this information system is subject to the constraints and peculiarities of the environment listed in II.5 and takes the name of MIS.

Defined as a "field of confrontation on its own" [dllea13], cyberspace is part of the informational environment, which comprises, according to North Atlantic Treaty Organization (NATO), the information itself as well as the individuals, organizations and systems that receive, process and transmit it, and the cognitive, virtual and physical space in which it occurs [Com12]. The following definition of II.3 is also proposed.

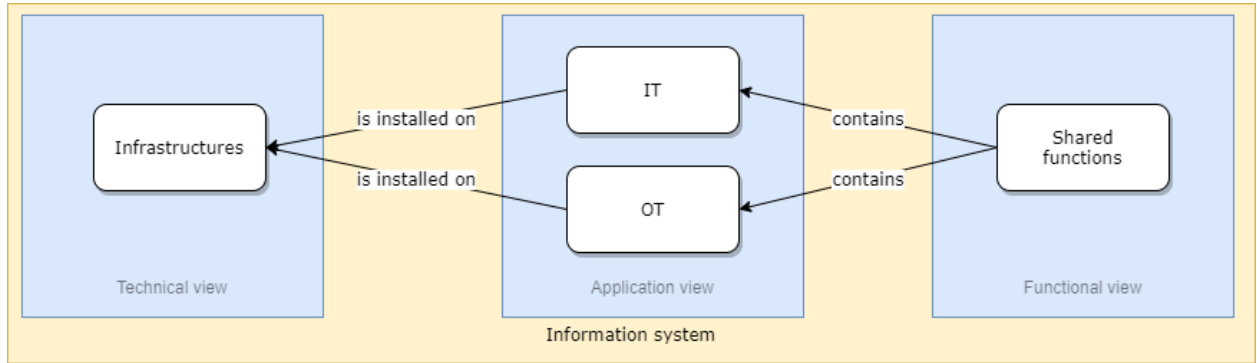


Figure II.6: Functional, application and technical views of an information system (source: personal work).

Definition II.3. *Cyberspace is the global domain made up of the mesh network of information technology infrastructures (including the Internet), telecommunications networks, computer systems, processors and integrated control mechanisms. [CIC]*

Cyberspace is interdependent on air, land, sea and space, areas of which it is however transverse, since it can be found in each of them⁹. If it is, thus, sometimes difficult to conceive a global representation, the American joint publication Joint Publication 3.12 proposes a modeling in three layers or zones¹⁰. First of all, the physical layer: it is the most concrete of the three zones, although recently made more diffuse by the techniques of outsourcing and *cloud computing*. It is made up of physical computer equipment and associated networks that ensure the storage, transport and processing of information. Then, the logical zone: it is made up of processes, tools and exchange flows, algorithms and protocols which are not based on a physical reality, but on a purely logical addressing. Finally, the social or cognitive layer, called *cyber-persona* in Joint Publication 3.12 is more difficult to understand. It abstracts information in digital format present in the logical layer to create digital representations of an actor or entity in cyberspace, including individuals and associated digital identities (avatars). In the maritime domain, cyberspace includes all the MISs mentioned in the previous section.

Definition II.4. *Cybersecurity is the desired state for an information system, allowing it to resist to events from cyberspace that may compromise the availability, integrity or confidentiality of data stored, processed or transmitted, and the related services that these systems offer or make accessible. [CIC]*

⁹We will also recall the etymology of the prefix *cyber*, taken from the Greek *Kubernêtikê* which means "rudder".

¹⁰Cyberspace operations of June 8, 2018, available at https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/jp3_12.pdf

Cybersecurity proposed by the definition II.4 is therefore the desired state for MISs to work in security. It is based on cyber protection, cyber defense capabilities and cyber resilience. It should be noted that no notion of voluntary action is indicated: events taking place in cyberspace can also have an involuntary or accidental origin. Environmental incidents (for example a fire), a hardware failure or a software anomaly involving information systems, often neglected in risk analyzes, must therefore be taken into account.

Definition II.5. *Cyber protection is the set of technical and non-technical protection measures allowing an information system to withstand events likely to compromise the availability, integrity or confidentiality of data stored, processed or transmitted, and related services that these systems offer or make accessible. [CIC]*

The definition II.5 of "cyber protection" makes it synonymous with "information systems security". In the maritime context, it is therefore the desired state for MISs to be safe and efficient in terms of availability, integrity and availability from their design and throughout their life cycle.

Definition II.6. *Cyber defense is the set of technical and non-technical measures enabling a State to defend in cyberspace the information systems deemed essential. [CIC]*

The suffix "defense" has led to some overuse of the term cyber defense defined in II.6 in the military context, to which it is however not reserved. The following specific definition II.7 is however given to military cyber defense.

Definition II.7. *Military cyber defense is the set of defensive or offensive actions carried out in cyberspace in preparation or in the planning and conduct of military operations, in particular to guarantee the effectiveness of the action of the armed forces and the proper functioning of the ministry. Military cyber defense complements network, system and information protection measures with a capability for operations in cyberspace and a cyber crisis management capability, coupled with the resilience capabilities of information systems. [CIC]*

In the maritime context, cyber defense therefore corresponds to the desired state for MISs to be technically capable of being defended against cybernetic aggression. Cyber defense also brings together a dedicated crisis organization, able to detect, qualify, counter a cyber event, or even initiate a digital operation.

The notion of cyber resilience is defined below (II.8):

Definition II.8. *Resilience is defined as the capacity of an organization to face events (incidents or aggression), to resist them and to recover. Applied to cyberspace, it is called cyber resilience and defined as the capacity of an information system to resist a failure and to return to its initial state after the incident. [CIC]*

In the maritime context, cyber resilience is the desired state for MISs to offer a high level of availability in the event of an attack, whether human, environmental, or simply following to a failure. Resilience is a daily requirement for sailors at sea, having to deal with unfortunate situations on their own: it logically corresponds to a characteristic they legitimately also expect from the MISs they use.

Cyber monitoring represents the real-time tool of cyber defense. It contributes directly to the "observe, orient, decide and act" (OODA) loop, because it offers a real-time capacity to detect any event that could prove to be hostile or unusual. A definition of cyber monitoring applied to MISs could be the one proposed in II.9:

Definition II.9. *Cyber monitoring is defined as the ability of an organization to detect any hostile or unusual event in an information system that could impact its availability, integrity, confidentiality or the traceability of its data. Applied to the maritime sector, it includes the installation of sensors on MISs, the collection of the metadata they produce and their analysis within a generic capacity for monitoring cyber defense in order to search, in real time. or deferred, for indicators of compromise. Cyber monitoring also comprises a dedicated organization and specialized tools that will allow the development of the reference cyber situation of MISs.*

II.4.2 Inventory of MISs vulnerabilities

Like any information system, MISs are affected by numerous vulnerabilities that could be exploited by an attacker [Jon16, Bur14, DiR15, Cim11]. However, three symptomatic difficulties can be considered as the main vectors of vulnerabilities.

First, the life cycle and patch management process of a ship: after an engineering phase and a construction phase of several years¹¹, ships are generally designed to sail for thirty to forty years. During this period, their use must be maximum in order to ensure profitability. This busy schedule complicates the major evolution of the ship's MISs. If patch management

¹¹Sometimes up to 15 years for the most complex military ships.

is possible on the office systems present on board and the IT in general, it is much more complicated to implement for the benefit of the OT. In addition, the lifespan of these systems is notably longer than that of the IT: it is thus realistic to think that certain systems will not be replaced at all - apart from major damage - during the lifespan of the ship. Once the ship is in operational status, the implementation of corrective measures, the securing of old protocols is unlikely, if not impossible at an acceptable cost. The eventual consequences of patch management on the operation of other complex ship systems have to be assessed, in a context of growing interdependence between systems.

Then, the maturity of the sector: maritime cyber security is a relatively new and until recently little discussed field, especially with regard to research. Even very recent research works on MISs, published during this thesis, make virtually no reference to maritime cyber security issues [LMWW21]. While the military issues of maritime cyber security have been assessed and taken into account overall over the past ten years, the civilian maritime world also seems to be less considered. The underlying issues are nevertheless essential. The many reasons that can explain the difficulty of taking into account the cyber security of MISs are: a weak perception of the stakes, a lack of governance and coordination, insufficient staff and crew awareness, a lack of knowledge of MISs, lack of policy from the shipowner or operator, complexity, cost, difficulty in identifying and retaining experts.

Finally, an almost generalized absence of cyber security equipment and particularly for the detection of cyber events is observed on board ships. As a result, detection can be late or even last for the life of a system until it is withdrawn from active service.

If these three difficulties lead to think that the sector is vulnerable, the observation must however be shaded. Indeed, digital technology is a vector of progress for the maritime environment, for the safety of people, the monitoring and safety of cargoes. From now on, it is no longer possible to imagine doing "without" digital. In addition, maritime sector standards frequently impose specific requirements in terms of resilience, especially for cyber physical installations. They thus generally provide for redundancy of these installations. They sometimes even have separate safety chains, operating in parallel with the control chains, to verify that the production operation is correct, in accordance with the expected values and within the limits set for the system. Finally, the role of man should be emphasized: the sailor is resilient, trained, experienced and able to endure difficult situations and react in the event of an unexpected event. This systemic analysis of the vulnerabilities of the maritime environment is summarized in the figure II.7.

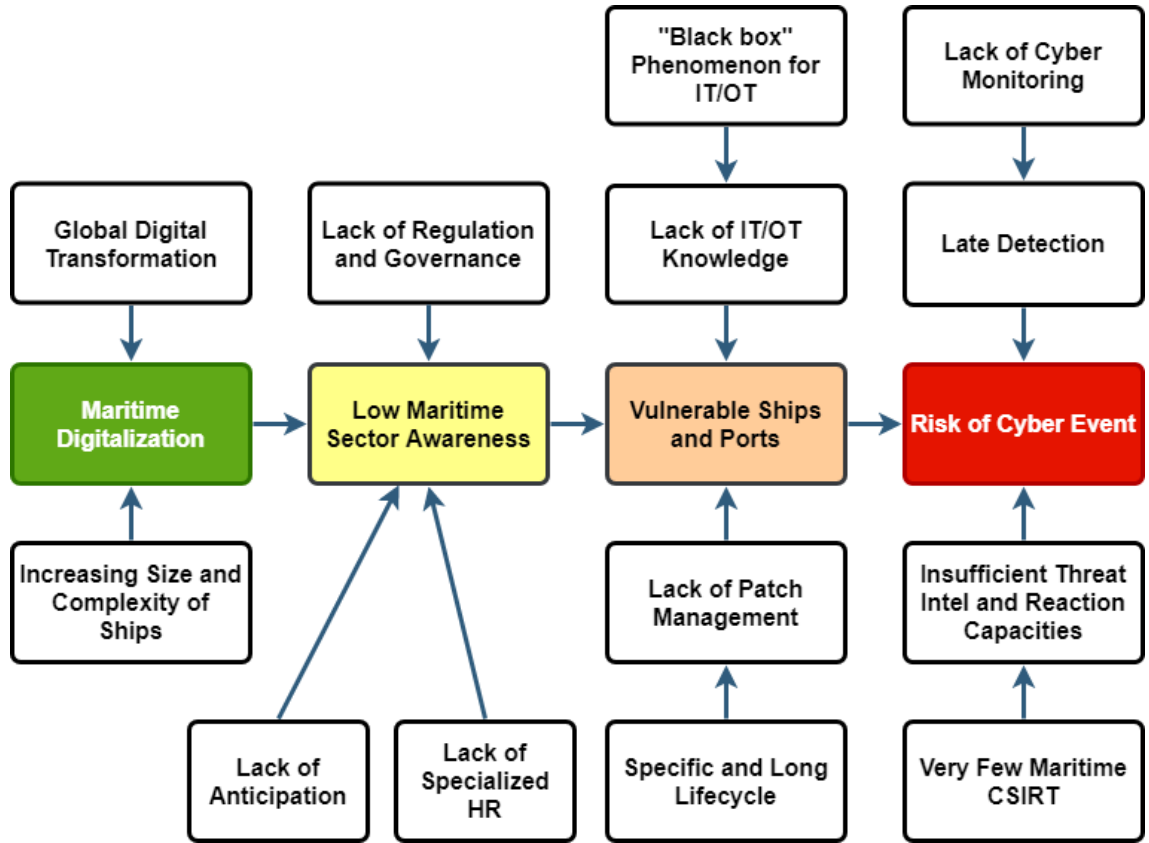


Figure II.7: Systemic analysis of cyber vulnerabilities in the maritime environment induced by digital transformation (source: personal work).

II.4.3 Exploitation of MISs vulnerabilities for offensive purposes

The advantage of using the cyber vector to conduct an offensive operation is well known: indeed, the attacker knows that his action, if it is carried out discreetly, is unlikely to be the object of reprisals (whether judiciary, cyber or kinetics), as the reaction is difficult to conduct in this confrontational field. First of all, the anomic nature of cyberspace makes it particularly hard to frame from a legal point of view, whether for truly legal or political reasons. Despite recent efforts in the field of international law [da19], it is indisputable that certain actors also have their interest to ensure that this state of affairs remains unchanged. Then, the rapid digital transformation of the sector is creating a moving cyberspace, in constant evolution, difficult to master and map and whose functional inter-dependencies are constantly increasing. These inter-dependencies could exist before the digital transformation: for example, the information system of a ship often requires that the servers and network equipment have an air-conditioned room. A cyber physical system provided this air conditioning function (chilled water production system, for example), but was generally

not digitized. However, digital transformation now affects industrial systems: the latter thus rely on an information system (touch screen, PLC) to be started, stopped and operated. The next step in digital transformation consists in increasing the interdependence of information systems between them (for example: dissemination of information relating to the availability of air conditioning, water temperature, etc.). Acting on a system in action or in response to an incident can thus induce consequences on other systems not initially impacted and essential to operations (collateral damage). Finally, the difficult accountability of actions in cyberspace facilitates manipulation and amplification operations. Manipulation consists of either producing, diverting or falsifying information for one's own benefit. Amplification aims to give a strong impact to an information, true or false, and that one would not have noticed otherwise. These two modes of action, frequently used on social networks, can take part in disinformation actions on a more strategic level. The reaction to this type of action makes it difficult to consider legal or kinetic actions, in particular because of the potentially global nature of the amplification, the short time frame of social networks, the necessary counterattack effort and the difficult attribution of the initial actors of these information warfare actions. Finally, the action is often based on evolving technical vectors which require high transversal skills: to guarantee its effectiveness, harmlessness and, in the case of an attack, the absence of long-term accountability is not not an easy thing.

There are many reasons to use cyberspace to target MISs [Fit15]. Notions of proof, accountability for actions, physical or political borders, distance, organization or time remain more complex to grasp than in physical spaces. Manipulations or amplifications also remain possible, which contributes to modifying the nature of the balance of power between adversaries and to reinforcing the asymmetry between attackers and defenders. The attacker often has several steps ahead, and he knows that the response to the attack will not be sufficiently dynamic, or even that the attack will remain undetected. To achieve his package, he can rely on several vectors, which can be used in combination. First, he can use the physical area: theft, trapping, physical destruction of all or part of the on-board MISs by using Electro Magnetic Pulse (EMP), environmental aggression (fire, water leak...). Then, the attacker can choose the logical zone: data theft (espionage), use of software vectors, network-centric attacks. Finally, it can also act in the social zone, by reaching the people or the families of people who evolve in the maritime domain, for example by exploiting Open Source INTelligence (OSINT).

II.4.4 Sources of threats, cyber risks for MISs

Even fragmented, the public data relating to cyber security events in the maritime world that we were able to identify highlight the predominance of sources of unintentional internal or external threats (crew members, external stakeholders, designers) on the one hand, and on the other hand from sources of voluntary external threats linked to cyber crime and generally not aimed strictly at the maritime sector, which therefore constitutes rather collateral damage in this case [Jac]. Attacks specifically targeting the maritime sector exist, but remain poorly documented publicly. Finally, it should be noted that the state actor threat is reinforced for the military component with also cases of data theft for espionage purposes, which can also target academics, manufacturers and subcontractors in the sector.

Cyber threats targeting the maritime industry can target information. In this case, the objective is then to exploit, extract, make unavailable, destroy, manipulate or corrupt the digital information: map data, navigation, information on the loading or transit of goods [Bea17], the passengers, the characteristics of the ship. Information systems can also represent targets. In this case, the attacks may be aimed at mapping the information systems, identifying their vulnerabilities, modifying or temporarily disrupting their operation, or even destroying them immediately or at an appropriate later time. Finally, physical cyber systems can constitute targets with high added value: the objective is then to use cyberspace as a vector to achieve physical objectives. It can thus act to disrupt, neutralize, or even destroy equipment, infrastructures or critical installations by acting on the computer systems which control or manage them by exploiting the computer vulnerabilities of some of their components. Threats can therefore target all three layers of cyberspace: physical, logical and cognitive. Annex B summarizes the risks for the sector.

II.5 Conclusion

As we have seen in this chapter, the maritime sector carries strategic challenges for French and global interests. It is of capital importance for many sectors of the economy as well as for the daily life of the French. In this rapidly changing sector, MISs play an essential role. Contributing directly to the proper functioning of ships and ports, they allow the safe performance of many functions, physical or otherwise. However, the particular characteristics of these systems expose them to different sources of threats. Intrinsically vulnerable, they can represent a major interest for the attacker, and the security tools available today are

poorly adapted and do not allow to perceive or understand the situation, nor to follow the evolution of a cyber event in time and even less to remedy it in constrained time. It therefore appears necessary to identify the tools and approaches most able to respond to the scientific, organizational, technological and human obstacles of maritime cyber security, to design those tools with an approach guided by the modeling of their operation and, finally, to verify the scientific and operational contribution.

*Maritime Cyber
Situational Awareness:*
definition, model and
contribution to
maritime cyber security

In this chapter, we discuss RQ1, the data collect architecture and its modeling necessary to improve detection, understanding and response to a cyber threat. We specify the contributions of the cyber monitoring of MIS to the detection of feared events in the sector and the difficulties relating to its implementation. We then present the current limitations of security products which constrain their direct use in the maritime environment. Finally, we detail the principle of *Situational Awareness* and propose its application to the cyber security of MIS through the original concept of *Maritime Cyber Situational Awareness*, for which we propose a functional architecture. The original work of the MCSA concept has been the subject of several international publications [[Jac18](#), [Jac19a](#), [Jac19b](#)].

III.1 Cyber monitoring of MISs: the problem

In this section, we first assess the interest of setting up cyber monitoring capabilities of MIS to respond to the structural difficulties encountered by the maritime world. Then, we study the current methods and tools that could meet the expectations of the sector in terms of cyber monitoring and identify their limits.

III.1.1 Contributions of cyber monitoring to the difficulties of the maritime world

Maritime cyber monitoring, as we defined it in II.9, aims to detect any hostile or unusual event, voluntary or involuntary, which could impact the MIS. It includes the organization, tools, human resources and methods of detection, classification of an event, and remediation. The cyber monitoring of MIS can rely on the organization, tools and procedures of an operational cyber monitoring center (Security Operations Center (SOC)), as it's the case in other sectors [Zim14]. The technical infrastructure of these centers includes judiciously placed and configured field sensors, subject to permanent local or remote monitoring. After data is collected by a local architecture, it is then sent ashore to a focal center, in charge of the analysis, detection and qualification of cyber events and having visualization and synthesis expertise. Several IT tools contribute to this mission, such as Security Information and Event Management (SIEM) and Cyber Threat Intelligence (CTI). If the border between SOC and Computer Security Incident Response Team (CSIRT) is the subject of recurring debates, in particular on the resolution of incidents, these centers also use Security Incident Response Platform (SIRP) in order to offer a capacity for dialogue and sharing with other cyber security players, to ensure crisis management, or to inform management levels to contribute to decision-making.

Apart from the detection, qualification and response to threats, cyber monitoring also offers a first response to the structural problems identified in figure II.7 and which we recall below. First of all, the monitoring provided by the CTI makes it possible to improve the anticipation of threats, whether it concerns vulnerabilities related to hardware or software, the fight against the viral threat or knowledge of Tactics, Techniques, Procedures (TTP) opponents [WDWI16]. Shortcomings in expert maritime cyber security human resources were recently highlighted by two studies, one English [Tam19] and the other carried out with French shipowners and ports [Ben18]. The concentration of these resources in an expert center therefore responds in part to the difficulties of preparing the sector in this area. Cyber monitoring, even on vulnerable and legacy systems, can detect attempts to exploit vulnerabilities - as well as their absence. If, in most cases, cyber monitoring architectures use "passive" devices and therefore do not allow a physical reaction, they nevertheless offer the technical, human and organizational possibility of detecting and reacting quickly. Without correcting the initial causes of the patch management problem, cyber monitoring therefore provides part of the answer. In addition, the ability of cyber monitoring to "listen" to hosts, servers and ship's networks in search of traces of compromise also results in a significant gain

in terms of maturity and knowledge of the IT/OT onboard or in ports. The tendency of the ship to constitute a "black box" is attenuated and the realization of dynamic maps of MIS becomes possible. Finally, cyber monitoring generates interesting side effects: for example, the finesse and volume of the information reported is of interest for predictive maintenance, the detection of configuration faults, but also for other purposes (logistics monitoring, mapping, mastery of operational capacities). The figure III.1 takes again the figure II.7 by underlining the added value of cyber monitoring in the treatment of vulnerabilities induced by the digital transformation of the maritime world: the positive contributions appear in blue. However, they should be qualified by recalling that, on the one hand, the means must continue to be mobilized to correct the systemic causes already mentioned and, on the other hand, that the implementation of cyber monitoring represents a significant cost and adds new equipment, systems and services that require long-term security, maintenance and administration.

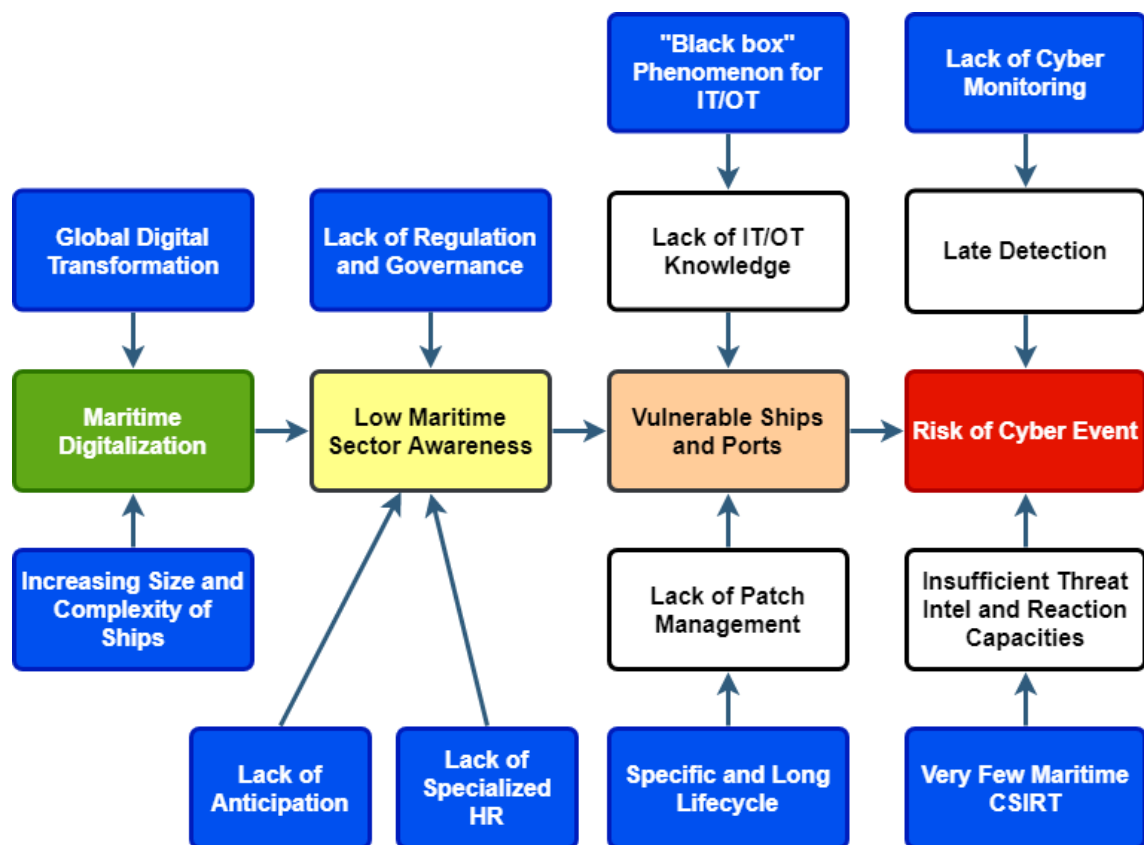


Figure III.1: Benefits of cyber monitoring for the maritime sector (source: personal work).

Table III.1: SOC needs *vs* constraints of the maritime world

SOC needs	Cyber maturity	Expert and trained HR	Remote intervention	Data volume	IT/OT knowledge	Patch management
Maritime sector	Relative immaturity	Limited HR	Complex intervention	Constrained connectivity	Black box effect	Lifecycle
Solutions	Expert center	Centralized HR	Logistics	Dedicated architecture	Mapping	Detection capacity

III.1.2 Inadequacy of classical cyber monitoring methods

The classic modeling of an organization and a technical architecture of cyber monitoring requires a variety of skills which come up against the current lack of maturity of the maritime community. Indeed, cyber monitoring must be based on existing means and procedures which, in the civilian maritime world, may turn out to be absent or immature. The human resources qualified to operate a maritime SOC are thus rare and take a long time to train. The intervention and the remote monitoring of a ship prove to be complex, due to the mentioned constraints of its connectivity, which also limit the capacity to transmit the volumes of data expected by the SIEM. Finally, the lack of knowledge of certain IT and especially OT systems makes it difficult to deploy sensors which, moreover, might not be suitable. The Table III.1 crosses the necessary capacities for the implementation of cyber monitoring capacities with the expected difficulties. If the establishment of a cyber monitoring capacity is complex, its contribution is real.

Designing a security model based on cyber monitoring also exposes new risks and requires attention to adverse techniques [Bra11]. An attacker, to hide his tactics, can attempt to saturate cyber monitoring tools with multiple detections which, for deception purposes, could camouflage a more complex and less visible targeted attack, the signal-to-noise ratio then proving unfavorable to the detection model. Among the unintentional risks, human resources not specialized or unfamiliar with the detection architectures, the monitored systems, or a bad positioning or settings of sensors can lead to a partial or distorted view of the real situation. Finally, since the sensor itself constitutes a computer system, it can also

be the subject of a compromise. It therefore appears necessary to design appropriate and secure technological mechanisms, as well as an architecture and compromise indices that allow surveillance to be contextualized. It is also advisable to test the functioning and the smoothness of these solutions by specific tests such as audit or intrusion test.

The market of cyber monitoring tools is proving to be rather mature and prolific. In particular, SIEM solutions are numerous and, for the most part, have been designed and improved for years. Network-based Intrusion Detection System (NIDS) or Host-based Intrusion Detection System (HIDS) type sensors and centralized logging software have also been around for a long time, especially on IT systems. The functions of a SIEM are essential for the detection and processing of a cyber event. Its role is to collect information, to aggregate it, to standardize it, to correlate it, to offer dashboard and alert capacities, to archive and to be able to replay events. However, these multitasking tools have limitations. Among these, Miloslavskaya [Mil19] underlines the difficulties of SIEM to work on a large scale of heterogeneous and geographically distributed systems, the difficulties of detection in the midst of multiple false positives and the consequent mass of data. The delays in detecting and processing alerts are often incompatible with the dazzling nature of a cyber attack. To fill these gaps, Cardena *et al* suggest the use of mass data processing technologies [Car13].

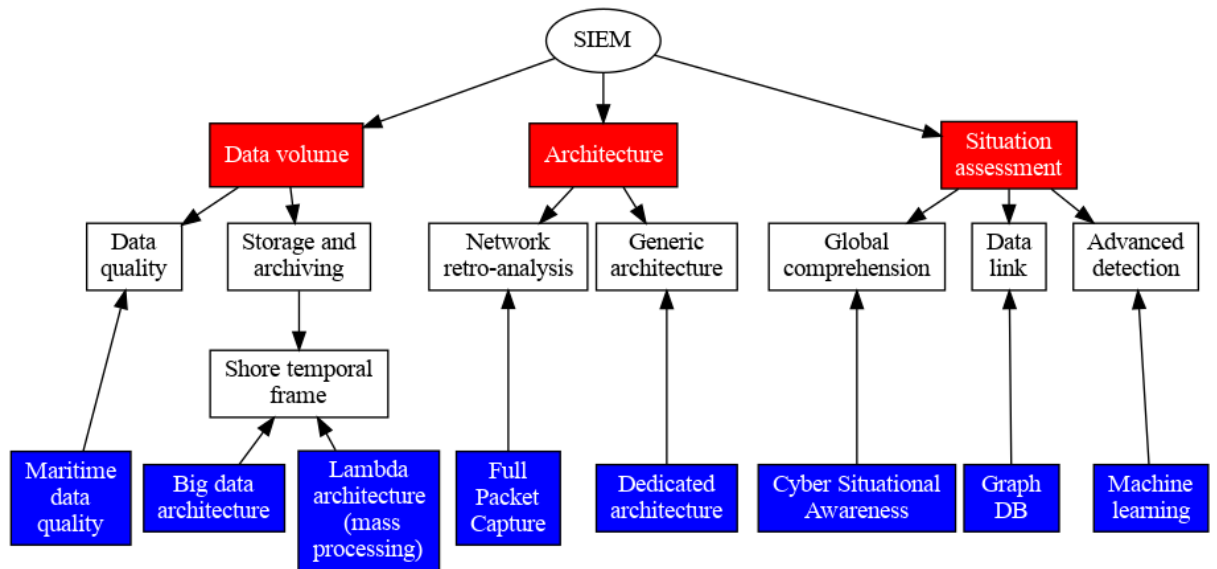


Figure III.2: Limitations of the use of classic SIEMs in the maritime context and possible solutions (source: personal work).

The figure III.2 models the current limitations of SIEM which can make their use

difficult in contexts such as the maritime sector, under aspects related to the volume of data (incompatible with the maritime environment), to architecture (lack of consideration of environmental constraints) and assessment of the situation (the tool in itself does not allow us to appreciate what is taking place). Faced with this problem, the possible solutions (represented in blue in the figure) will mainly be the subject of studies within the framework of this thesis.

III.2 Situational awareness

During discussions carried out within the framework of our research work with SOC operators, we identified that their main difficulty is linked to the lack of overall understanding of the information from SIEM, facing a large number of alerts and a - relative - knowledge of the multiple systems monitored. These difficulties of perception, comprehension, projection and resolution encountered in the context of cyber monitoring allowed us to establish a parallel with the concept of Situational Awareness (SA).

III.2.1 General context

A situation represents the characteristic of one or more agents confronted with events, circumstances and relations between objects (figure III.3) [Fer97]. They must be able to position themselves and assess the situation in order to make decisions.

The definitions, interpretations and applications of SA are numerous, whether the origin is scientific, military or industrial [Tad10]. We have retained the definition III.1 which seems to refer and which is proposed by the researcher Mica Endsley, at the origin of the concept.

Definition III.1. *Situational Awareness is the perception of elements of an environment in a volume of time and space, the understanding of their meaning, and the projection of their state into the near future. [End95].*

SA therefore brings, in a volume of time and space, the notions of perception, understanding and projection into the near future of a situation, with the objective of human decision-making with a view to action in relation to a situation.

NATO recommends the use of the term situational awareness and gives the following definition III.2:

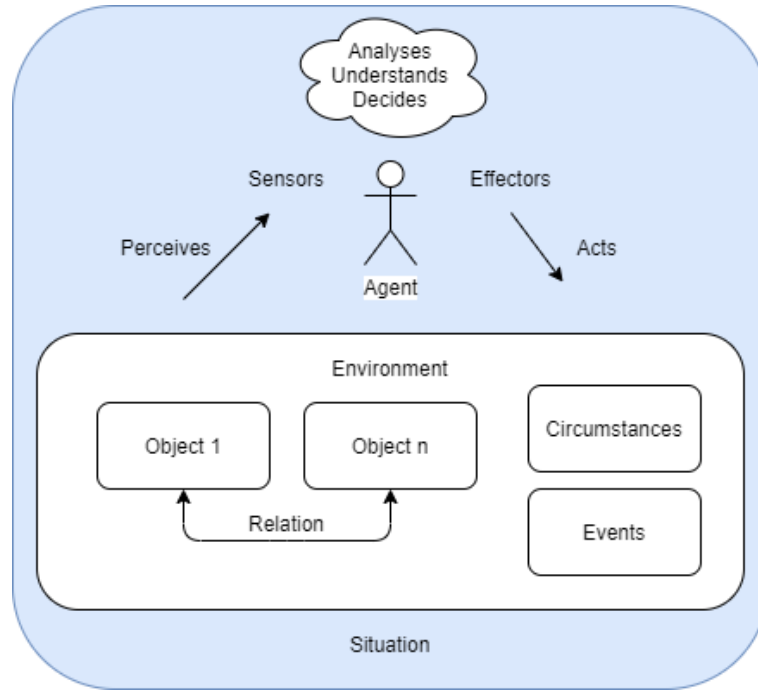


Figure III.3: A man, as an agent, in his environment and facing a situation (source: personal work).

Definition III.2. *Knowledge of the elements of the battlespace necessary to make decisions based on appropriate information. [OTA17]*

In France, the Ministry of the Armed Forces speaks of "situation understanding" and gives the following definition (III.3):

Definition III.3. *In the areas of strategic anticipation such as the conduct of an operation, level of understanding resulting from the identification of trends and links that develop over time, in space, between actors and aimed at making their comparison with the observed situation. [CIC]*

It is therefore a *state of knowledge*, in Endsley's own words, while the organizational or technical processes of acquisition, processing and analysis are grouped together under the term "Situation Assessment". They directly contribute to the elaboration of SA, but are more of a process *per se* rather than a true cognitive objective. In a military context, the French Ministry of the Armed Forces gives the following definition (III.4):

Definition III.4. *A process of logical reasoning which enables the leader to take into consideration all the factors influencing the military situation and to arrive at a decision concerning*

*the conduct to be adopted in order to accomplish his mission*¹.*[CIC]*

The process of acquiring, processing and analyzing data (Situation Assessment) is based on two large families of data, namely that concerning our own situation (Knowledge of Us (KU)) on the one hand, providing sufficient knowledge of one's own state (position, operating state, etc.) and, on the other hand, gathered data relating to others (position in time and space, knowledge, skills, intentions...), grouped together under the term of Knowledge of Them (KT).

We can understand the state of knowledge brought by the SA in response to three questions, of increasing complexity and level of maturity: for Situation Perception, it is the answer to the question "What are the current facts?", for Situation Comprehension, this is the answer to the question "What is happening?", for Situation Projection, this is the answer to the question "What is most likely to happen if?". Endsley thus establishes a 3-level model (Layer 1 SA, Layer 2 SA, Layer 3 SA). To the state of knowledge provided by these layers, Mc Guinness & Foy suggest adding a fourth level, called Situation Resolution, which aims to be aware of the resolution possibilities and therefore facilitate decision making by answering the question "What exactly should I do?"*[McG00]*. This fourth level constitutes an essentially human assessment today, which is based on processes - generally technical - of a lower level, which can be standardized and automated and which themselves rely on information from sensors. This decision support must also propose options for the decision maker by having identified, for each of them, the possible impacts. The desired final state of the SA consists in allowing an informational (information superiority, or even information dominance) and decisional (decision superiority) superiority in the face of a situation. The figure III.4 constitutes a synthesis of the whole.

Finally, it seems appropriate to note that, on complex systems of systems, the global state of knowledge provided by the SA can be composed of several more "thematic" SAs. On the other hand, to achieve a global understanding, Franke and Brynielsson underline that the integration of the thematic SA into the global SA is essential *[Fra14]*. The global SA can be seen as a merging of intermediate SAs. However, just as an intermediate SA is a state of knowledge, so is the global SA. Concretely, it should make it possible, for example, to detect a cause and effect relationship between the knowledge of a degraded cyber situation and consecutive or simultaneous knowledge of a degraded situation of an access control system or a cyber physical equipment. However, some significant pitfalls remain on the subject

¹In the maritime military framework, the mission could be considered as a special circumstance (for example: projection of a combat ship on a maritime zone to accomplish a mission).

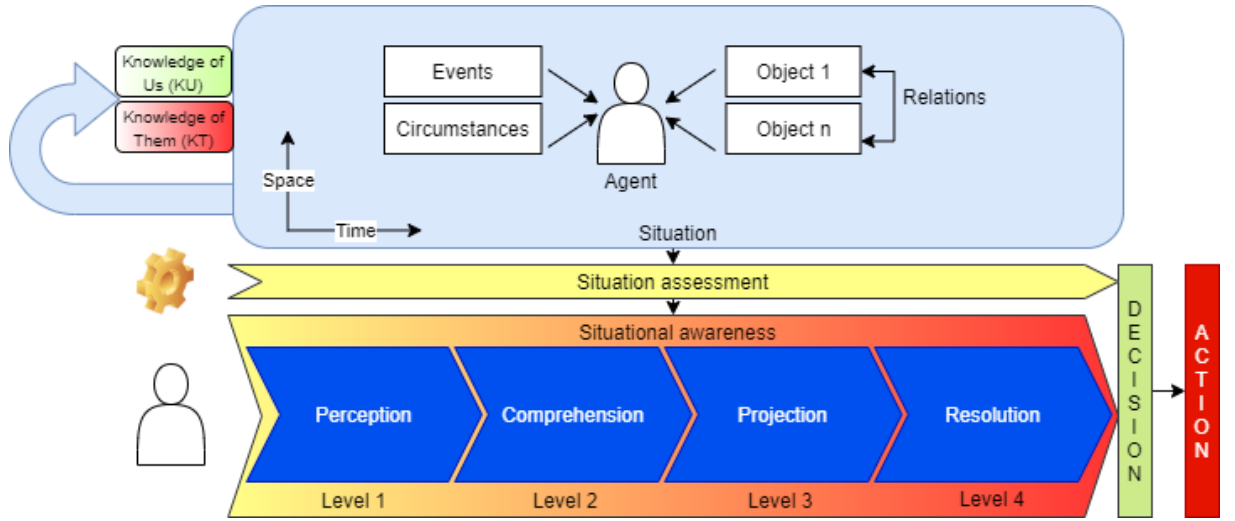


Figure III.4: Situational Awareness elaboration modeling according to Endsley [End95], adapted with the Situation Resolution component suggested by Mc Guinness & Foy [McG00] (source: personal work).

of the SA merger. Indeed, without sufficient abstraction, the risk of being submerged by the quantity of information obtained from intermediate SAs cannot be neglected because no tendency or understanding can be detected to lighten up the understanding and therefore knowledge. Then, the interest of a global SA lies in the existence, the richness and the quality of the intermediate SAs: it therefore appears essential to put them in perspective.

A difficult point for achieving the fusion of intermediate SAs lies in the fact that they do not necessarily exist or, at best, that their compatibility is not assured. Indeed, no norm or standard of exchange today makes it possible to integrate the various SAs and any attempt in this direction thus requires a particularly important effort. To achieve this global SAs, at the output of Endsley's model for a system, the existence of a normalization step before the export of the intermediate SAs would be necessary. In the first stages of the process of constituting the global SA (perception and comprehension), the intermediate SAs must also be able to be correlated with each other, most probably by the through high scenarios based on feared events, or by a more geographical approach if the events occur, for example, on the same site, or even in the same room.

III.2.2 First application of the SA in an aviation context

Developed primarily for use in the military air context by Mica Endsley, the goal of SA is to function effectively in peace and crisis time. The Head Up Display (HUD) (figure III.5), is

Table III.2: Examples of KU and KT data for air context SA elaboration

KU examples	KT examples
Environmental values	Enemy RADAR detection
Sensors and systems malfunction	Enemy electronic warfare detection
Position, navigation, time	Enemy missile detection

regularly suggested as the perfect example of the information that the pilot must have at his disposal to react correctly and promptly when faced with a situation. As a real aid for in flight decision-making, the instrument does not replace the pilot or his environment, but interfaces between the two. For example, in peacetime, the head-up display gives the pilot information relating to his current and future position in space. It thus has essential information on a surface of a few tens of square centimeters and possibly some additional screens [Foy05]. Most of the information displayed comes from physical sensors (heading, speed, altitude, angle of attack, engine power, external pressure, etc...) and logical information (route to follow, availability of on-board installations, etc...). All knowledge relating to one's own environment is referred to as KU [Tad10]. The head-up display also displays information relating to other nearby aircraft, hostile or not, and their evolution in relation to the pilot's aircraft. Finally, for military aircraft, onboard sensors benefit from frequent updates to improve enemy detection capabilities using electronic warfare equipment. The detection of an enemy aircraft or radar, the launch of a missile and their evolution in time and space can be grouped under the term KT. Table III.2 synthesizes some examples of KU/KT data for the elaboration of the aerial SA.

The success of the HUD led to its generalization to all modern combat aircraft (where speed of decision-making is essential in a fast kinematic context), and more recently to commercial airliners, and even to terrestrial vehicles. New functionalities have been added to the HUD, such as the display of information from the Traffic Collision Avoidance System (TCAS) protocol, which assists in the prevention of the collision risk between two aircraft. The resolution is achieved by ordering opposite maneuvers to the two aircraft [Kuc07]. However, the role of man is never neglected: the pilot and the radar control centers remain responsible for detecting aircraft in their areas of responsibility, but also for making the final decision. Communication channels also exist between aircraft and control centers to allow information sharing, and therefore the development of the final air SA on a larger scale using Automatic Dependent Surveillance-Broadcast (ADS-B)² and Aircraft Communications and Reporting

²This protocol has several similarities with AIS, including in terms of vulnerabilities to a software defined radio threat [BSG⁺20].



Figure III.5: Head Up Display of a Bombardier CRJ-900 passenger jet in long final at Munich airport (source: [Wikimedia Commons](#), unmodified picture, author: Joschiki, license: [CC BY-SA 3.0](#)).

System (ACARS).

III.2.3 Application of SA to the maritime context

A major objective of the actors of the maritime world is also to achieve the state of knowledge provided by the SA, sometimes without being fully aware of it. Data from a ship's sensors are generally centralized before being merged and presented to the watch officer at the bridge. On recent vessels, a large number of screens are thus placed there to display KU-type information, such as the current and projected position and kinematics (calculated from the speed over ground, the heading, the angular velocity, drift...), orders for the engine or fire/flooding alarms. These multi-function screens can also display KT-type information. It shows the position of nearby ships, for example by using AIS. Some screens also show the fusion of KU and KT information. This is the case for the "surface SA", where the vessel's own position is merged, on the same screen, with the AIS information from other vessels

Table III.3: Examples of KU and KT data for maritime context SA elaboration

KU examples	KT examples
Environmental values	Enemy RADAR detection
Sensors and systems malfunction	Enemy electronic warfare detection
Position, navigation, time	Enemy missile detection
Sensors and actuators values	<i>Automatic Identification System</i>
Sonar and RADAR data	Intelligence flow

and other echoes. Similarly with the air context of the TCAS protocol, information from AIS and navigation radar is of major interest in detecting the risk of collision with other ships (Closest Point of Approach (CPA)). Identical solutions will allow the development of the SA for autonomous surface vessels from remote centers [Por14]. Table III.3 summarizes some examples of KU and KT data for the elaboration of the civil or military maritime SA.

As in the air context, on military vessels, the sensors and associated computing systems on board are updated to improve the detection and identification capabilities of enemy threats. Ships can also exchange information with each other and with land-based centers, to enable information sharing and the development of an extended or global maritime SA. Unlike in the aviation context, the integration constraints on board a ship are generally less, the volume available for the installation of sensors or servers being greater and the certification aspects less restrictive in most cases. This partly explains the presence of many screens and the lesser need to merge SA in the bridge or, for the moment, the lack of recourse to HUD or augmented reality. Apart from the standards proposed by the National Marine Electronics Association (NMEA) [A⁺02, All], no protocol or standard exists for the sharing of information between devices. As a result, SA is not achieved through one or two screens and a head-up display, but often through a significant amount of screens. On recent ships, multi function screens (Multi-Functions Display (MFD)) provide a possible transition between several visualizations which allow to obtain alternately different SA, each visualization also being able to display and carry out control/command options (orders to the machine, opening and closing of mechanisms, acknowledgment of alarms, *etc.*).

On board a ship and in ports and surveillance organisations, tools such as Vessel Traffic Service (VTS) on land, or Electronic Chart System (ECS) and ECDIS on board allow the perception and understanding of the maritime situation from a decision-making perspective. [KHL⁺21] Modern VTS (*cf* figure III.6) and ECDIS (*cf* figure III.7) make it possible to elaborate a maritime tactical situation by merging information obtained from various sensors and databases and their display. This is also the case with combat systems (Combat

Management System (CMS)). They replace, on the one hand, the use of a paper chart for the follow-up of the ship's own position and its projection and, on the other hand and at least partially, the radar screen for the monitoring of the surface situation. They also allow a certain form of projection: estimation of the future position, (calculation of the Estimated Time of Arrival (ETA), for example) and possibly the calculation of collision prevention alarms (CPA). Human decision support (*situation resolution*) is still not very present in the current versions, but is the subject of numerous announcements for the years to come from the manufacturers of maritime electronics.



Figure III.6: US Coast Guards VTS operator (source: [Wikimedia Commons](#), unmodified picture, author: US Coast Guard, license: public domain).

As previously mentioned, on a ship, several "thematic" SA can coexist, such as the physical SA (protection of the ship), the fire SA, the maritime surface SA, the air SA or the cyber SA, without any merge of these different SA being effective today, mainly due to the lack of protocol and interconnection standards, but also because the merge could make understanding and visualization complex. A European project, SAURON, is carrying out research on the fusion of different SA for the purpose of hypervision³. As a summary, if

³<https://sauronproject.eu/publication.php?id=28>



Figure III.7: Example of an ECDIS Human Machine Interface (source : [Wikimedia Commons](#), unmodified picture, author: [Hervé Cozanet](#), license : [CC BY-SA 3.0](#)).

the captain of a ship has a good vision of his immediate maritime environment and of the coherence of the kinematics of his ship, and of the operation of his installations, he currently has no automated merge of the different SAs, nor the integration of cyber defense into its SA.

III.2.4 Operational processes for *Cyber Situational Awareness*

Cyber Situational Awareness (CSA) is the application of Endsley's model to the specific challenges and context of cyberspace. This variation has enabled several researchers to propose improvements or clarifications compared to the initial model, but which remain largely valid also for other use cases. Thus, Barford *et al* [Bar10] established that the CSA is recognized as effective when seven processes are developed to lead to the state of SA. The first process consists in establishing the knowledge of the current situation (situation

perception, P1), mainly from data collected by cyber sensors. The second process is to assess the impact of the attack on information systems and the operational processes and the evolution of this impact if no action is decided (impact assessment, P2). In order to achieve this, the process is based on knowledge of one's own vulnerabilities (vulnerability assessment) and threat assessment. Often carried out in parallel with other processes, the knowledge of the evolution of the situation over time is an essential management function (situation tracking, P3). The aim of understanding the attacker's behavior (attacker's perception, P4) is the process needed to determine the probable evolution of the attack (projection situation, P7), by crossing the knowledge of the attacker (KT) with the vulnerabilities of the systems to be defended (KU) and by determining several plausible scenarios depending on the attacker's likely intentions, opportunities and capabilities. In order to avoid the occurrence of the event and to seek an imputation, it also appears essential to have knowledge of the origin of the attack (why, how) (causality assessment, P5). Finally, the quality level of the elements that constitute the CSA (quality assessment, P6) must be continuously assessed, to ensure the relevance of Endsley's model and the conclusions resulting from the conduct of the various processes. This point can be considered as part of the process of perception.

The seven processes appear to be complementary: the development of a comprehensive and quality CSA requires that they all be carried out and evaluated. A CSA in which one of the elements is missing would become incomplete. The processes are also interdependent: a process of approximate quality would have an impact on the quality of the others and of the whole. Thus, the overall quality of the CSA depends on the P6 process which sets quality criteria for the inputs and outputs of each process: CSA's overall quality level therefore results from the individual quality of each process. The processes are essentially developed from parameters or information obtained from the monitored information systems or depending on sensor quality criteria. They can also be a function of the quality of exogenous factors. Thus, the development of certain processes requires recourse to information from external sources (for example, CTI). The assessment of these criteria is essentially based on humans: while the elements can and must be acquired, transmitted, processed and displayed by the machine, their final assessment (consistency, inconsistency, quality) still often appeals to humans. Finally, as shown in III.8, maturity and complexity increase from left to right and from bottom to top of the model: thus, it is more complex and more successful to determine the future evolution of the situation than to detect a current threat. Decision-making is also a complex process, as it involves both man and machine. It also does not seem conceivable to try to establish a projection of the situation without first realizing it. Finally, it should be noted that the development of the CSA is carried out by analyzing KU and KT elements

known *a priori* (mapping, compromise indices, scenarios) in relation to KU and KT elements *in vivo* (observables).

Figure III.8 provides a visualization of Barford's model of the constituent processes of CSA. As part of our work, we have adapted this model with respect to [Bar10], so that the projection process is not linked to the level of understanding, but also in the perspective of adding two processes. Indeed, Barford's model does not incorporate [McG00] proposal to add the situation-solving process. We therefore propose the addition of two associated processes at this level. The first, anticipation resolution, aims to anticipate the resolution of a cyber situation by writing scenarios based on best practices, but also by capitalizing on incidents that have already occurred (playbook). The second process, resolution assessment, has the objective of evaluating the best resolution scenario and measuring its possible impacts on the information system and business processes. Finally, SOC's and CSIRT's are sometimes interconnected and can exchange real-time information on incident detection and processing using standards such as Structured Threat Information eXpression (STIX) and Trusted Automated eXchange of Indicator Information (TAXII): sharing and merging to the higher or inter-SOC and inter-CSIRT levels are therefore essential functions. SOC's and CSIRT's must also report to the SA at leadership and command levels, with a consistent level of abstraction.

The CSA concept remains an active research topic. Peng Liu thus analyzed the main research questions associated with CSA (figure III.9), to which we could also add those relating to the resolution of the situation presented in the same article [Bar10]. Our work has made it possible to contribute to these subjects, with the exception of risk and doubt reduction, hypothesis management and machine learning.

By relying on data from sensors, the situation assessment process with a view to obtaining the CSA must provide automated tools or human interactions allowing decision-making in order to react to a (voluntary or involuntary) cyber event [Onw12]. In the cyber context, KU information is essentially composed of real-time flows directed to a SIEM and sometimes associated with databases (asset management or information systems vulnerability database). KT information consists primarily of data from the CTI, either through the use of Indicator of Compromise (IoC), or directly through the use of consumer or private detection signatures. IoCs can thus be made up of file names or hashes, domain names, IP or email addresses. But the CTI also provides other valuable information for the SA, such as TTPs, which detail the habits and techniques of attacking groups. Table III.4 summarizes some elements of KU/KT for the cyber context.

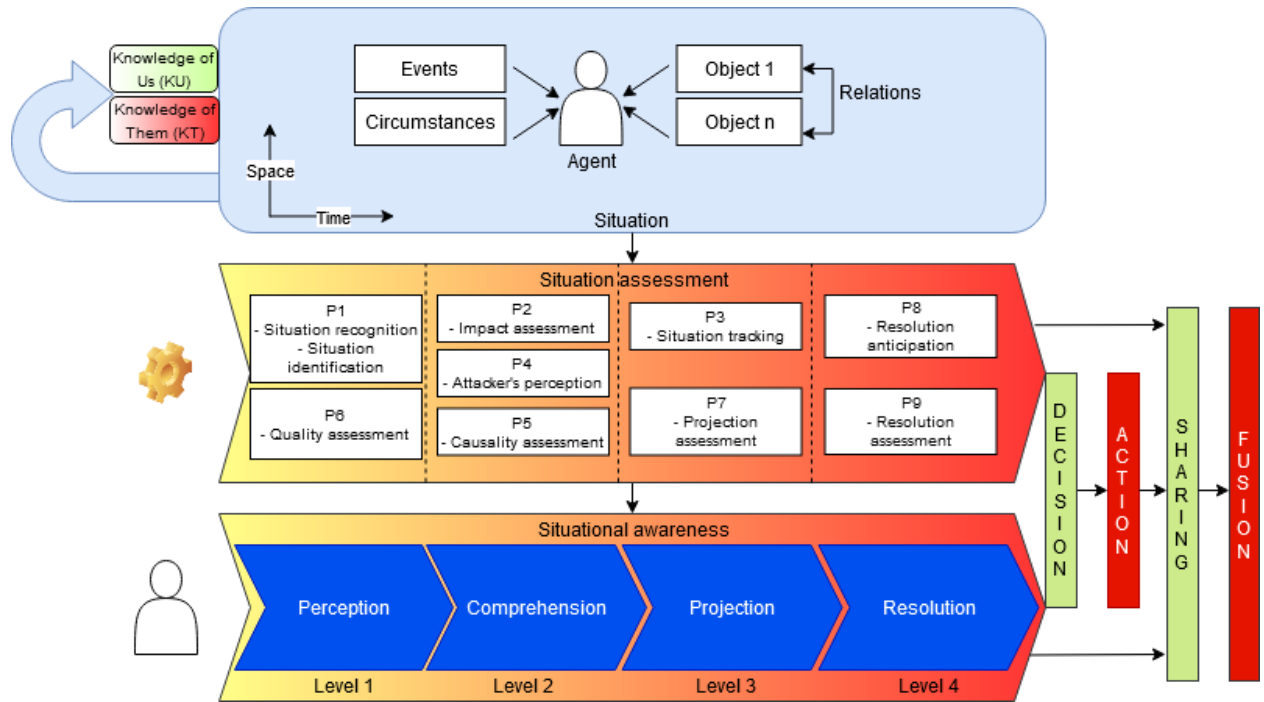


Figure III.8: Constituent processes of the CSA, work adapted from Barford and Mc Guinness & Foy [McG00] and taking into account the remarks of Franke and Brynielsson [Fra14] (source: personal work).

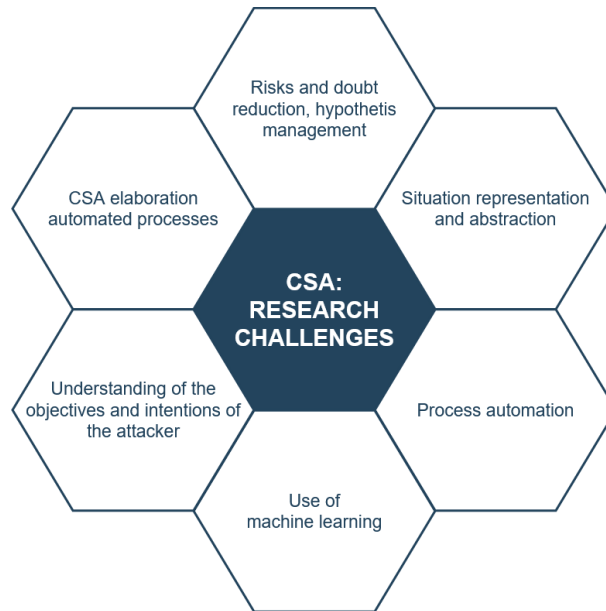


Figure III.9: Main research subjects related to CSA, according to Peng Liu [Bar10] (source: personal work).

After having analyzed the three contexts (air, sea, cyber) and the corresponding

Table III.4: Examples of KU and KT data for CSA elaboration

KU examples	KT examples
Digital assets knowledge	IoCs from CTI : IP and email addresses, filenames, files hashes, JA3, domain names. . .
Network flows knowledge	Sensors signatures (public or private)
Metadata from sensors	TTPs, realtime detections

Table III.5: Examples of similarities of KU and KT data for the elaboration of air, sea, cyber SA

	Air	Sea	Cyber
Real-time KU	Sensors, calculators, HUD and screens	Sensors, calculators and screens	Metadata flows from cyber sensors (NIDS, HIDS...)
Real-time KT	Radars, IFF, electronic warfare	Radars, electronic warfare, sonars	Alerts from cyber sensors and IoCs
Threat-based sensors updates	Mainly military sensors (targets updates, signatures)	Mainly military sensors (enemy radars signatures...)	Onboard cyber sensors (NIDS/HIDS) updates with public or private signatures or IoCs
SA sharing	TCAS, ADS-B, ACARS, radio	AIS, radio and others	STIX/TAXII

KU/KT, we propose to make a synthesis (Table III.5). While the sensors behind the KT/KU differ, the variety of sensors and heterogeneity of data are similar. Exchanging information and abstraction to higher levels are also essential capabilities.

As mentioned earlier, a research debate persists on the CSA sharing method. Franke and Brynielsson recall that the CSA does not constitute a separate SA, but that it must be part of a global SA [Fra14]. However, the current organization of SOC now often consists of a distinct organization from other decision-making centers where "sectoral" SAs are merged [Zim14]. Therefore, when existing, the global SA scarcely integrates the cyber SA.

III.3 Modeling the *Maritime Cyber Situational Awareness*

In this section, we discuss the modeling of the operational/business and functional processes that allow the development of the MCSA. Like Figure III.8, this operational/business process modeling should identify the sensors necessary for the development of the MCSA that will appear in the technical architecture. The modeling of the functional architecture should, for its part, make it possible to identify the different information, fusion and processing flows necessary for the development of the MCSA. At each stage, the specificities of the maritime sector identified in Figure II.5 must be correctly taken into account.

III.3.1 Example of state of knowledge for a proper MCSA

As a state of knowledge, the MCSA can be seen as the sum of elementary knowledge projected into time and space (real and virtual) specific to the maritime sector. Definition III.5 provides an example of the state of knowledge that could be acquired following the development of the MCSA:

Definition III.5. *The maritime information system [name of the MIS] has been the subject of a cyberattack since [hour]. The attack has for origin [origin] and its objective is [objective]. The impacts are [confidentiality-impact, integrity-impact, availability-impact]. The cause of the attack is [identified_cause]. The attacker has [attacker's_visibility, attackers'_means]. The timeline of the attack is [timeline]. The cyberattack could continue with [attack_projection]. To ensure the continuity and resumption of activity and return to a safe nominal situation, short-term and long-term measures [solutions_resolution] should be taken and [global_SA] informed.*

III.3.2 Cyberattack modeling: the example of the ATT&CK model from MITRE

While several possibilities exist for defining the different modes of cyber attacks, they do not, however, all have the same objectives or the same scope [Mav17, Al-16]. Based on this observation and the need to better detail the attacker's operating methods, the MITRE organization has developed the MITRE ATT&CK™ (Adversarial Tactics, Techniques, and

Common Knowledge) framework, which formalizes and models the various attacks that can target a information system, as well as a cyber analysis directory⁴ [Str17]. This framework makes the assumptions that, on the one hand, the defender has significant knowledge about the attacker's TTPs and that, on the other hand, the detection of threats by an antivirus is unreliable to cope with Advanced Persistent Threat (APT)-type threats. Indeed, the attacker, by using obfuscation technics, can check the detection capacities of the market antiviruses to the viral charge he designs. The concept therefore assumes that (organisational or technical) classical defenses will be bypassed and it focuses on behavioral analysis, using a threat-based and regularly updated model.

Several critiques can be posed on this framework in its current version and definition. We can formulate five main ones. First, with respect to the life cycle of a cyber attack, the MITRE model is logically positioned after the exploitation phase. It does not seek to ensure detection of the phases of reconnaissance, weapon preparation, delivery of the charge or the exploit itself: the detection model is therefore intended to be advanced, but late in the cyber kill chain process [Yad15]⁵. Second, to be effective, the framework must be applied across the entire information system of the company and have the associated and configured sensors and centralization, in order for example to detect lateral displacement. The information system must have a high level of maturity in terms of cybersecurity⁶. Third, due to its post-exploitation positioning, it mainly focuses on the analysis of system values by sensors to ensure detection, to the detriment of network analysis, for instance ([Str17], §1.2). Fourth, if it appears appropriate for the operating systems present on the information systems of a company, it is difficult to apply to a part of the Industrial Control System (ICS), especially with regard to PLC, because the model is focused on systems for general public operations such as MicrosoftTM WindowsTM or GNU/Linux. Fifth, the ATT&CKTM framework today focuses on identifying legitimate functions (for example, system functions like *schtasks.exe* under Microsoft Windows) that could be exploited for cyber attack purposes, but do not rely on existing IoCs that could be disseminated through the CTI. These limitations therefore slow down its application on MISs, or in any case establish that the model alone is insufficient: the recognition step prior to a cyber attack remains likely to generate a lot of noise and false positives at the level of the NIDS sensors and make it difficult to detect advanced threat as

⁴See <https://car.mitre.org/>

⁵Since then, MITRE has also developed a framework model on the early cyber attack phases, called MITRE PRE-ATT&CK MatrixTM : <https://attack.mitre.org/matrices/pre/>.

⁶This means, for instance, that the organization should have a fully-mapped IT, cyber sensors or a SOC. See also ISO/IEC 21827:2008 and the cyber maturity guide of the French national information agency, ANSSI.

sought by the ATT&CK™ framework. In contrast, the early detection of a threat is more effective if the recognition step is included in the detection process. The use of the framework alone will allow detection of more advanced threats, but at a later stage of realization, thus hampering capacities for action; it should therefore be considered as one step among a set of detection measures that rely on generic, but advanced, models of attacks. Figures C.1 and C.2 in appendix C model the example of an attack on an ECDIS-type maritime system in the pre and post-exploitation phases, as well as the countermeasures that could allow detection.

The identification of these technical countermeasure means must be considered as an essential step to be carried out in parallel with the identification of the exploitation phases. As such, HIDS-type intrusion detection sensors provide essential cyber defense capacities. They offer the possibility of carrying out cryptographic integrity checks of configuration or binary files, the collection of system events and event logs, compliance analysis or even the software inventory of client and server workstations. While they have the advantage of allowing a detailed analysis of the configuration and the integrity of a host, they must however be precisely adapted to the configuration and the software installed. For example, HIDS sensors will need to be configured to scan the ECDIS installation and configuration directories, otherwise it will not be able to detect a change there. NIDS-type sensors are historic cybersecurity tools that remain fairly widely deployed. They can be exploited for many purposes, such as signature-based or behavioral anomaly detection. They also allow anomaly detection on temporal or contextual variations in network flows (sudden increase compared to usual bandwidth, etc.). However, they remain particularly sensitive to false positives [Das14] and poorly suited to networks where the use of cryptography would be important. However, their interest is real in certain phases of the preparation or execution of the attack, as well as for protocol dissection for the purpose of metadata feedback. The ideal positioning of these sensors depends intimately on the architectures to be monitored: in the context of ships and port infrastructure, MISs can turn out to be very varied and complex. In general, HIDS agents must be deployed on each client and server station. Regarding the NIDS, their positioning should be considered as essential between two interconnected information systems, between the ship and shore assets and, if possible, listening to each network loop, in particular on OT (ECS/ECDIS, ICS, etc.). The fusion, correlation and analysis of data from the two types of sensors are therefore essential.

III.3.3 MCSA abstraction levels

The military world generally organizes operations according to three layers, which make it possible to define levels of abstraction [CIC]. The strategic level is the highest: it is "the level of political leadership and command at which a state or a group of states sets national or multinational security objectives and provides national, inter-ministerial, especially military, resources for their deployment in different theaters". The operational level is "the higher level of military command projected into a theater or responsible for a geographical area, to which operations are planned, conducted and supported by one or more forces in order to achieve the military objectives set by the strategic command and thus contribute to achieve the desired final effect". Finally, the tactical level is "the subordinate level of military command projected onto a theater or an area of operation, at which actions are planned and executed to achieve military theater objectives and contribute to the achievement of the desired end effect with means assigned to the operation within each component". In a simplified way, we can consider that, at the level of the Ministry of the Armed Forces, the General Staff of the Armies constitutes the military strategic level, the Maritime Zone Commander the operational level and the commander of the ship or the constituted force (carrier strike group, amphibious group) the tactical level. This type of abstraction in three layers can also be applied to the civil sector, where a similar division could be made: the management of the shipowner would constitute the strategic level, the operations manager by maritime area or type of vessel the operational level and the captain of the ship the tactical level.

However, other levels of abstraction must also be analyzed in the dual context of cyber and maritime. The first level is the level of cybersecurity skills. On board the ship, the tactical level does not generally benefit from a cyber expert: he must therefore have a level of information abstraction that allows him to understand the situation without going into too much technical detail. The captain of a civilian ship must therefore know if he can continue his mission (impact assessment), if a degraded mode is available and the actions he can take at his level (resolution assessment). At the operational level, cyber experts are generally present (for example, within a SOC). They are the ones who will determine many elements of the situation assessment and the modes of action. At the strategic level, the management does not always have high expertise in cyber, but needs, like the captain of the ship, to measure the impact on its operations and to identify if the incident that has occurred could also have an impact on other MISs. A third possibility of abstraction is based on a geography or mission oriented aspect: the operational level can be linked to the deployment

Table III.6: Needs in terms of technical data, abstraction and global SA integration for the three different levels of MCSA

	Technical need	Abstraction need	Global SA integration need
Strategic level	Low to moderate	High	High
Operative level	High	Low to moderate	Low
Tactical level	Low to moderate	Moderate	High

zone of a ship: thus, an operational manager can be designated for a given maritime zone (for example: Atlantic, Mediterranean, Pacific). A fourth possibility is to consider only one type of vessel in particular. Finally, a fifth level of abstraction must be able to be carried out for each functional type of MIS (evaluation by domain). For example, it may prove relevant to have a state of knowledge on the cyber situation of all the systems ECS/ECDIS of a fleet (navigation domain) or of all the installations of propulsion of ships (propulsion field). Each elementary MCSA must also be able to transmit its elements to the global SA (by ship or by domain).

The table III.6 summarizes all of these elements.

The selected modeling must therefore make it possible to integrate these five dimensions (by organizational level of abstraction, by skill level, by geographical area or mission, by type of vessel, by functional area of MISs), but also to be able to merge the MCSA in the global SA. Large-scale analysis of multi-dimensional heterogeneous layered networks, sometimes referred to as relational networks, requires identifying models that are simple, intuitive, representative and allow the transformation of inter-network relationships into efficient graphs to perform calculations. Several models exist, in particular models based on non-oriented hypergraphs [Yen10], as represented in figure III.10, where the non-oriented hypergraph (H) is composed of a set of vertices (V) and of a set of hyper-edges (E). The hypergraph makes it possible to represent a set of interconnected subgraphs.

Modeling using super networks makes it possible to represent the inter-dependencies between different graphs [Jgu16, Zha19] and constitutes an interesting alternative. The interest compared to hypergraphs resides in particular in the inter-layer connection, which can be achieved through particular nodes, in a layered rather than hierarchical mode of operation. The super-networks also allow a better representation and visualisation of the

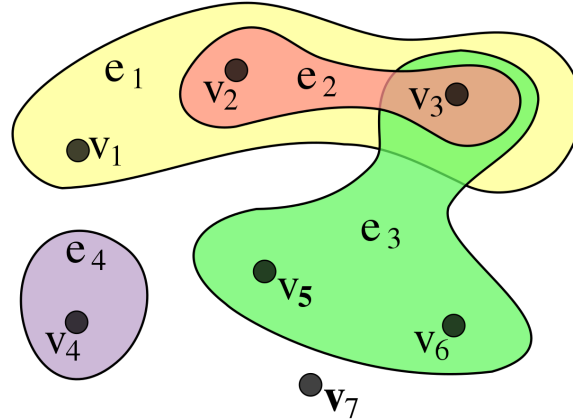


Figure III.10: Example of a hypergraph (source: [Wikimedia Commons](#), unmodified picture, author: [Kilom691](#), license: [CC BY-SA 3.0](#)).

different functional layers which, in our case, should allow us to facilitate the taking into account of the abstraction levels mentioned. The figure III.11, using a modeling based on super-networks, thus makes it possible to represent the five different levels of abstraction identified, as well as the possible inter-dependencies between the information systems. This modeling could also make it possible to integrate the notions of horizontal and vertical propagation of the confidence [Cos18], essential to the establishment of a MCSA of quality.

III.3.4 MIS cyber monitoring system data flow aligned with SA operational processes

The objective of this subsection is to detail the processes necessary for the elaboration of the *situation assessment* in the cyber context, based on the processes identified in figure III.8. This work is the first essential step in the development of the SA. To our knowledge, no scientific work available has focused on the identification and characterization of these data flows in the maritime context. The numbering of the processes resumes that proposed in figure III.8. These processes are intended to be generic to the establishment of a CSA and not specific, in the current state, to the consideration of environmental constraints for the establishment of the MCSA.

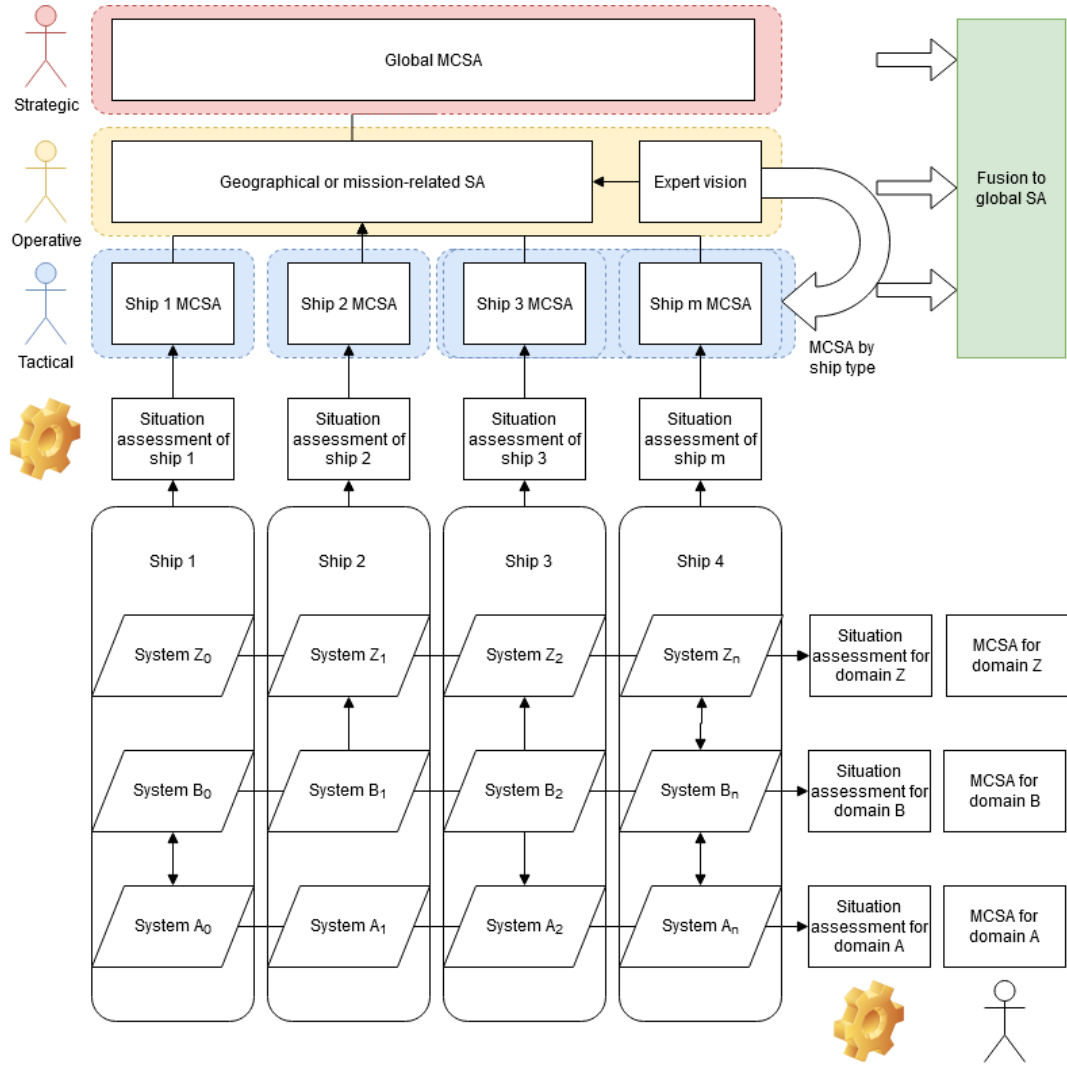


Figure III.11: MCSA modeling using super networks (source: personal work).

III.3.4.1 P1 - *Situation recognition and identification*

This process must make it possible to detect and identify the existence and type of cyberattack (category, precise description), the source (who, what) and the target (who, what) [Dut13]. This process is the first detection process. In near real time (the main case of our work) or based on preventive or reactive investigations for isolated or unmonitored systems, one of the main difficulties of this process is linked to the level of detection: whether the system is based on signatures or behaviors, the risk of false alarms in the face of a very large amount of data collected is a real issue. First of all, the background noise (made up of false positives (FPs)), due to poorly suited signatures, can be so significant that analysts can be quickly overwhelmed. This technique can also prove effective for an adversary to

drown the SOC under a deluge of detections, in order to saturate the analysis to mask a more advanced attack. In this context, the attacker also has an advantage over the defender, as many NIDS signature sets are made up of public data (for example: *Emerging Threats*). Then, the type of sensor, its intrinsic capabilities (ability to analyze specific protocol and extract metadata), its configuration and the signatures it uses can generate false negatives (FNs). This error rate (we also speak of precision by calculating $\frac{TP}{TP+FP}$) is to be compared with the rate of true positives (TPs) and of true negatives (TNs) to determine the quality of the whole (we can also calculate the recall $\frac{TP}{TP+FN}$). This first process, the only one linked to the direct capture of data on MIS, is essential: without a good quality of detection, it is all the processes of *situation assessment* and the *situational awareness* which will be altered, potentially causing real difficulties in understanding the situation and in decision-making (decision-maker myopia).

After capture and normalization, the comparison between KU and KT (metadata generated on a MIS *versus* IoC of an attack) or KU and KU (for example, value or metadata generated at a time t_1 *versus* its value or its reference cardinality at a time t_0) allow to perform the anomaly detection. Table III.7 summarizes the type of possible comparison.

To be effective, this process therefore requires the presence of sensors (NIDS, HIDS, producers of security logs: antivirus, firewall, syslog...), pre-processors adapted to the protocols used and a good knowledge of the monitored systems. The choice and positioning of these sensors in the MIS architecture are essential and may depend on the type of system being monitored and its network architecture. Thus, the HIDS will be positioned, when possible, on all the computer workstations and servers of the ship's MIS. The positioning of NIDS can be more complicated on a ship when there are many subnets. Ideally, a NIDS must be positioned on each industrial network loop, on each specific network (navigation, communication...) as well as at each interconnection between two networks (for example between information systems bridge and those related to propulsion, in case of the presence of an Integrated Bridge System (IBS)).

It should also be noted that one of the difficulties in detecting for NIDS stems from the fact that they cannot always analyze the sometimes specific network protocols used on board ships (absence of a dissector for specific industrial or maritime protocols), and thus not have the advanced signatures that would provide a high level of precision.

Table III.7: Comparisons for detection

Objective	Example
Comparison of cryptographic hashes	Comparison of integrity between KU items issued from a HIDS with the NSRL database ⁷ . The hash can also be issued by a NIDS in case of unencrypted files transfers.
IoC comparison	Comparison of KU data issued from a HIDS or a NIDS with KT data issued from CTI (IP address, domain name, email address, etc...) and known to have been employed in recent attacks.
Traffic and signatures comparison	Comparison between network traffic (clear or encrypted, depending on the protocol layer on which the detection phase is conducted) between KU data (network transfers) and KT (private or public signatures).
Comparison between logs and Sigma-like databases ⁸	Comparison between KU data issued by HIDS or logging processes and KT-like expressions renowned to be regularly exploited during attacks (private or public signatures).
Metrics comparison	Comparison of metrics tendencies compared to KU data. The objective is to identify if measured metrics differ from a normal behaviour (KU). For instance: volume of data exchanged on a period of time or depending of time, used protocols, used TCP/UDP ports, <i>etc.</i> In order to limit the number of false positives, it can be helpful to use the cardinal characteristics of certain values in order to detect any new entry.
Machine learning	Comparison of data by supervised or unsupervised machine learning.

III.3.4.2 P2 - *Impact assessment*

The impact measurement is generally based on the existence of a physical, logical and functional mapping (knowledge part) and of a Business Continuity Plan (BCP)/Disaster Recovery Plan (DRP) or a crisis management organization to react to an attack. The impact measurement is also closely linked to the P4 (attacker's perception) and P7 (projection assessment) processes and to the knowledge of the possible vulnerabilities of MIS. Indeed, the measurement of the impact carried out at a t time can only be based on the elements available at the end of the process P1. Cyber incident forensics can be time consuming and understanding the attack mode and end goal may be delayed. Thus, the part of the MIS hit on that date may differ from the attacker's final goal and only represent an intermediate phase. This process, also called Impact Assessment or Battle Damage Assessment [OTA17], or even Cyber Battle Damage Assessment [BUTN18], to use a military term, includes analysis the current impact (damage analysis), but also the projection of the future impact of the attack if it continues or if the attacker has the capacity to exploit new vulnerabilities.

III.3.4.3 P3 - *Situation tracking*

The situation monitoring process is generally based on a software tool such as SIRP. The SIRP can be fed directly from the data coming from the P1 process (after correlation, in order to avoid the creation of multiple incidents) and play the role of scheduler towards analysis tools. The main difficulty lies in the treatment and analysis of multiple incidents, which may appear isolated, but in fact depend on the same incident, or on the same overall action. This process makes it possible to follow up, step by step, the situation, with its evolution, the decision-making, the movements of the attacker and all the events. This process is essential for decision making. Indeed, it makes it possible to group together under a single "case" a multitude of related events, but diffuse, in time or in cyberspace. However, it should be remembered that the attacker, if he knows he is being watched, can also use deception and denial of service techniques. The attacker will create multiple false alarms in order to spread confusion in the analysis teams and to slow down or divert the decision-making and the attention of the real target. Regarding denial of service, the attacker can decide to saturate the detection and monitoring systems with multiple very verbose alerts, in order to hide his real attack in the middle of hundreds of alerts which, poorly or not prioritized, will distract the defenders from the actual target.

III.3.4.4 P4 - *Attacker's perception*

Knowing the attacker is based on knowing the type of attack (tool, flaw) and TTPs. An attacker can indeed repeatedly use the same infection vectors or the same procedures to exploit the Cyber Kill Chain. Knowing your method allows you to anticipate the next stage of the attack, even if the attacks can also benefit from an opportunistic nature or from chance (collateral damage). This process makes it possible to know the behavior of the attacker, starting from the evolution of his known attacks and the projection of his intentions. For that, it appears essential to know the TTPs, generally obtained by the means of CTI. It continues with the *Projection assessment* process (P7).

III.3.4.5 P5 - *Causality Assessment*

The process of evaluating the cause is based on the analysis of logs and malicious codes, but also by the field investigation and forensics with the possible projection of experts at the place of the event. This process makes it possible to seek out and identify the origin of the attack (how), the reason for the attack (why) and, sometimes, the who (or at least an intermediary in the chain of responsibilities). This criterion includes causality analysis (going back in time to identify the various artefacts precursors of the incident) and digital investigation (inforensics). The analysis of a cyberattack and the origin of its cause can take days or more, which makes it difficult to assess this process in real time. As a result, it is usually dealt with on the sidelines of the SA establishment. If the attack vector can sometimes be identified quickly (for example: USB medium), the precise cause, the identity of the attacker or his motivations can thus take longer to establish, or even remain undetermined.

III.3.4.6 P6 - *Quality Assessment*

In a multi-sensor context, the quality of information is essential. From its contextual (measurement), intrinsic (sensor) and extrinsic (network) quality will depend directly, by propagation, the overall quality obtained during the fusion of the information [ML17, Bar10, Tad10]. Our work is based on existing sensors (NIDS, HIDS, etc.) which allow an analysis of the network or the system. We have therefore decided to concentrate most of our work on intrinsic quality measurements. This qualitative process, carried out throughout the *situation assessment* process, should make it possible to check the quality of the input

elements (sensors, KU, KT), before they are merged and then retained for decision-making. As part of our work, we propose the selection and monitoring of four intrinsic dimension criteria for NIDS type sensors. We have chosen to detail them, first of all because they are quantifiable indicators that would validate our model, but also because they remain insufficiently described today.

The first quality criterion, Q1, aims to check the reliability of the sensor (trustworthiness). Indeed, the sensors must remain available, be integrated, have up-to-date signatures and the link with the SOC must be ensured.

The second quality criterion, Q2, aims to verify the veracity and solidity of the input elements (veracity/truthfulness or solidity/soundness). In some cases, particularly in the case of detection, the mechanisms used are based on signatures. These signatures, if incorrectly written, can result in false positives and, if found to be incomplete, false negatives. As is often the case in anomaly detection, the verification of this criterion is based on data sets, sometimes old, representative of certain attacks. The detection of zero day attacks is therefore not possible using these signatures. In addition, it is often a question of data coming mostly from network sensors (for example, data recorded in .pcap format), and not enough from multi-sensor data (networks, event logs, etc.).

The third quality criterion, Q3, must verify the exhaustiveness of the input elements (completeness). In fact, the sensors, in particular the network, can ignore certain input packets, which can hamper their ability to, for example, subsequently reconstruct complete files. The elongation link and satellite bandwidth constraints represent all factors of potential data loss. The objective is therefore to measure Q3.1, corresponding to the loss rate of the NIDS, Q3.2, the loss rate due to the satellite link and Q3.3, the variation in the number of packets captured over time.

The fourth quality criterion, Q4, checks the freshness of the input data. It is the difference between the capture timestamp and the index timestamp in the SIEM. We also associate the measurement of timeliness (time scheduling of metadata on its arrival ashore).

These four criteria should allow us to evaluate the quality of the information coming from the NIDS sensors within the framework of the development of the MCSA. They will be the subject of a more formal definition during the experimentation on a maritime platform.

III.3.4.7 P7 - *Projection assessment*

This process is necessary for the projection into the future of an event in progress, in order to determine the next paths that the adversary could take, based on his perception and on the TTPs. This process is therefore acquired, on the one hand, by analyzing the attacker's intentions, his capacities and a factor of opportunism and, on the other hand, by the detailed knowledge of the possible attack paths established by a precise mapping of vulnerabilities.

III.3.4.8 P8 - *Resolution anticipation*

The objective of this process is to facilitate decision-making by anticipating the various possible incident response scenarios, also called Course of Action (CoA). Faced with each type of threat and potential target on the information system, this process makes it possible to establish scenarios which, planned and formalized, facilitate and accelerate the response to an incident by proposing one or more options for action. This formalization can be the subject of generic defense plans, and continuity and recovery plans by system. It should be noted that, due to the levels of abstraction identified in the table III.6, this realization is necessary at different levels (strategic, operational, tactical), actions, priorities and modes of execution may indeed vary from one level to another. The improvement and the permanent consolidation of this process and the associated CoA are essential to face a very diverse threat such as cyber. We therefore propose to enrich it, based on experience feedback, depending on internal incidents, the evolution of TTP or even feedback from external incidents.

III.3.4.9 P9 - *Resolution assessment*

This process is based, on the one hand, on the nature of the attack, the source, the target and, on the other hand, on the resolution anticipation scenarios of the P8 process. The environment (circumstances, events) can, however, influence this decision-making process. Indeed, the position of the ship, its situation (in port, at sea, in areas of operations), its kinematics, the missions or actions that it must carry out in the near or long time can weigh on the decision-making. Other contextual elements will also come into play, such as the existence of other means in the area (for example, if a ship can carry out the mission of the one affected by the attack), deployable personnel to come for mitigation or the possibility of investigation or action at a distance. Finally, any decision has consequences: thus, the

isolation of a ship's satellite link, the shutdown of information systems or even the use of backups have direct or indirect effects that must be measured, evaluated, discussed and taken into account during the decision-making process.

It is therefore by relying on the action scenarios proposed by the P8 process, on the contextual elements indicated and on the potential consequences on the mission of the ship and the owner that the decision will finally be taken (and not on exclusively technical or cyber bases).

It should also be noted that it appears essential to verify that this decision-making is followed by effects and that its realization, at all the levels of abstraction mentioned, is effective and efficient in that it must interrupt the attack and allow the restoration of the integrity, availability and confidentiality of the systems: this part of P9 can also be compared to the P2 process (impact assessment).

III.4 Benefits of visualisation for the KU of MISs and MCSA

The visualization of security events is a major issue in cybersecurity. Numerous research works, books and scientific conferences⁹ exist on the subject. In France, recently, several theses and research works have been published in connection with the subject [Hum15, KDC⁺18]. The research challenges remain important and the difficulties linked to visualization remain complex [Tad10]. Indeed, in the case of cyberspace, even if we rely on physical equipment, the representation that we can make of it often remains conceptual and abstract: it cannot in fact generally be produced in a geographic information system, taking into account moving mobiles adding additional complexity. The integration of the temporal dimension in the visualization of the situation is also essential, in order to allow a return to past data if necessary: the volume of information is therefore particularly substantial and its visualization over these important periods of time proves to be hard. Visualizing the projection of data into the future is also a real issue, which remains possible for continuous data, but more difficult to conceptualize for abstract data, such as the mobility and intentions of an attacker in an information system. Taking into account the specificities of cyber physical systems also remains a research issue and the visualization of their operation is

⁹In particular the IEEE Symposium on Visualization for Cyber Security <https://vizsec.org> conference

of undeniable interest [Sic18]. Finally, the fusion of information and its optimal presentation at different levels of abstraction are imperative to obtain a relevant and efficient MCSA.

Moreover, one of the systemic causes of the difficulties in setting up effective cybersecurity in the maritime world is explained by the lack of knowledge of these systems (*cf* figure II.7). The volume of data collected, which depends in particular on the type of sensor and its configuration, must therefore provide a richness and finesse capable of reducing this systemic cause and provide a better knowledge of the MISs. Thus, it appears essential that the work carried out within the framework of this thesis focuses on two essential goals for mapping a MIS: the analysis of the network, in order to understand the exchanges within the MIS and between MISs, and the analysis of the components of this MIS to identify its constituents. These two analysis are closely linked to the capacities of the sensors with regard to the interpretation of the multiple protocols, sometimes proprietary, present on a network. These analyses can be carried out in a completely passive way (which is generally the case for cyber monitoring where there should be, by design, no interaction between the sensor and the analyzed network), or active, in the case for example of a computer, where an agent must generally be installed (or remote queries carried out) to collect the necessary information.

III.5 Conclusion

In this chapter, we have mainly approached RQ1 and the subject of the assessment of the situation, as well as the difficulty of implementing an effective cyber monitoring without means adapted to the particular context of MISs. We have detailed the limitations of current cybersecurity tools. We then modeled the original concept of Maritime Cyber Situational Awareness and detailed the expectations in functional terms of this state of knowledge. We have adapted the original work of Endsley [End95] by enriching it with the work of Barford, Mc Guinness & Foy [McG00] and taking into account the remarks of Franke and Brynielsson [Fra14]. We have proposed the addition of two new processes (resolution anticipation and resolution assessment) which facilitate decision making. We modeled a cyber attack targeting a MIS using the frameworks proposed by MITRE, specifying the sensors to use to detect threats. Subsequently, we defined realistic and effective levels of abstraction in a cyber and maritime context, based on the theory of super-networks. Finally, we have detailed each process of the situation assessment, a level of detail absent in the state of the art of the CSA and we have approached RQ2 and RQ3 through issues related to visualization in a

maritime context. This work made it possible to show the important expectations of the scientific and maritime world on the subject, but also to identify a certain number of new lines of research, such as the automation of decision support and the measurement of impacts [Jac18, Jac19a, Jac19b].

*Maritime Cyber
Situational Awareness:*
architecture,
experimentation and
results

This chapter deals with RQ1 (collection and modeling architecture), RQ2 (analysis and representation of data) and RQ3 (anomaly detection in a maritime context), by applying the modeling and measurements of quality presented in the previous chapter. After having defined, implemented and evaluated the cyber monitoring architecture necessary to allow the elaboration of the Maritime Cyber Situational Awareness, we detail the results obtained in terms of quality, but also of data analysis and representation. Finally, we are directly contributing to an advanced detection architecture in the context of MISs, and more specifically on the NMEA 0183 standard.

IV.1 Design and experimentation of a platform prototype

The purpose of this section is to detail the platform experimentation carried out as part of our work, both on the on-board component and on the elongation towards the shore. The realization of this prototype will make it possible to verify the technical feasibility of our model for the elaboration of the MCSA.



Figure IV.1: Overview of three loops of the experimentation platform (industrial systems part) (source: personal work).

First, we detail the functional architecture of a modern ship, and in particular the layer enabling information to be captured for cyber purposes. Secondly, we specify the shore component installed on our prototype.

IV.1.1 Onboard component

The French Naval Academy has acquired equipment representative of maritime information systems, financed within the framework of a recent European and regional project. In addition, the academy recently acquired additional equipment within the framework of the European H2020 *Foresight* project, in order to extend its capacities to represent the operation of a ship's bridge. To model the platform prototype so that the whole platform is as close as possible, in terms of realism, of a ship, we based our work on three scientific articles which tackle the subject [Kri13, Rø14, TO21], as well as on articles concerning more particularly industrial systems [Nic19]. The prototype of this generic ship, designed in the form of a platform called *Naval Cyber Range*, should allow us to validate the technical feasibility of the elaboration of the MCSA (figure IV.1).



Figure IV.2: Simulation environment of the navigation bridge (source: [Naval cyber defense chair](#), reproduced with the kind permission of the author).

The full platform architecture can be represented according to several approaches. From a synthetic point of view, however, if we are interested exclusively in OT-type systems, the simulated ship can be considered to consist essentially of three subsystems.

The "bridge" subsystem, visible in figures IV.2 and IV.3, is representative of a medium-sized ship's bridge, with mainly GNSS and AIS sensors and associated antennas, a bus and connectors to NMEA0183 and 2000 standards, ECS display screens and a 3D navigation simulator. This subsystem is represented at the bottom left of the figure IV.5.

On board a ship, many industrial productions are carried out: electricity, propulsion, refrigeration, filtration. For these purposes, shipyards and their subcontractors create industrial installations that are ultimately quite close to what can be found in ports or other onshore installations, in the form of Industrial Control Systems, within the meaning of the definition proposed by the NIST¹. The "industrial" sub-system comprises four sub-networks, using network loop technologies: the "mobility" loop, the "security" loop, the "electricity" loop and the "auxiliary" loop. The "mobility" loop (figure IV.4) manages the main and auxiliary propulsion, as well as the steering gear: stabilization, rudder, azimuth thruster and shaft line. The "security" loop manages the detection and the different sensors and equipment to fight against disasters (water leakage, fire), the opening and closing of valves, doors

¹<https://csrc.nist.gov/publications/detail/sp/800-82/rev-2/final>



Figure IV.3: GPS receiver (left), NMEA 2000 bus (center) and AIS receiver (right) (source: personal work).

and panels, ventilation and the production of cold water. The "electricity" loop controls the production of high voltage energy on board the ship, according to voltage and current requirements. From sources such as turbines, generators or even batteries, electricity is then converted and distributed on board. Finally, the "auxiliaries" loop controls the management of the fuel, air and oil supply, the production of fresh water, the treatment of gray and black water, the refrigeration of food supplies, *etc.*. Within each loop, these functions are performed by one or more PLCs, to which are connected sensors and actuators which can be real or simulated. This subsystem is represented at the bottom right of the figure IV.5.

The "communication" subsystem represents a satellite communication between the ship and shore assets, with its constraints (bandwidth, delay, temporary loss of link). This subsystem is shown schematically at the top left of the figure IV.5.

The prototype experimental platform presented in figure IV.5 turns out to be quite similar to other models, such as the Purdue model for industrial systems and cyber test beds [TXLJ19]. It is intended to be as generic as possible, because each type of vessel has specific features in its sensors, actuators and systems, which can therefore be added in the process and instrumentation layers. In addition, the network architecture can also prove to be more complex, in terms of redundancy or physical or logical separation.

In order to carry out the experimentation and validation of our research models and

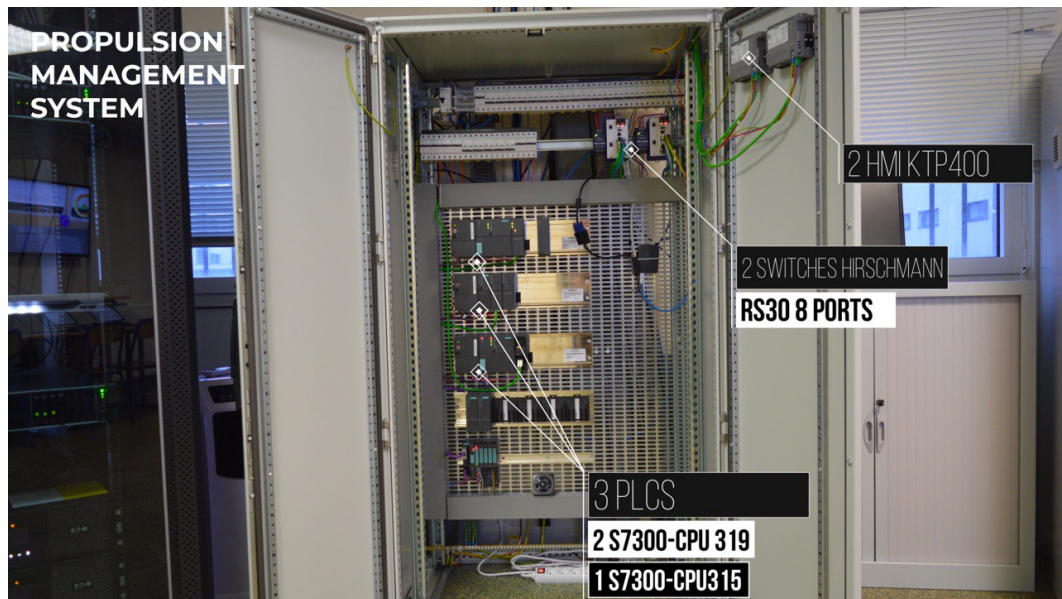


Figure IV.4: Overview of the mobility loop (the NIDS sensor is at the upper right angle of the rack) (source: [Naval cyber defense chair](#), reproduced with the kind permission of the author).

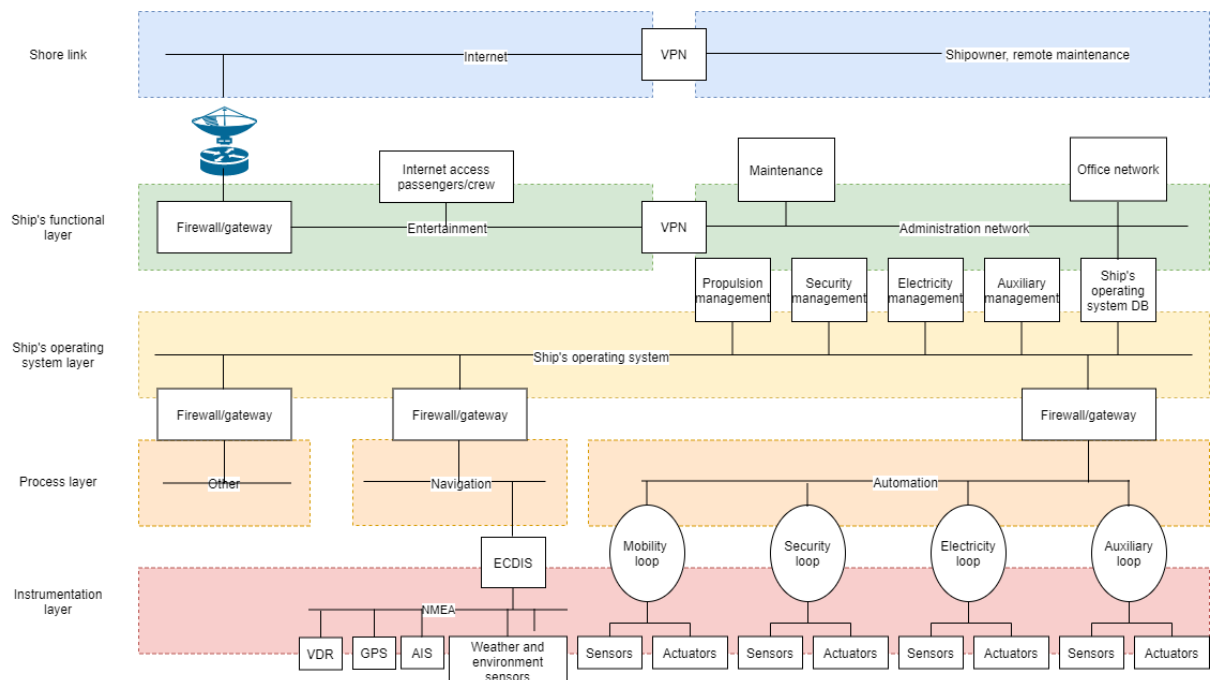


Figure IV.5: Technical modeling of the prototype with the "bridge", "communication" and "industrial" subsystems (source: personal work).

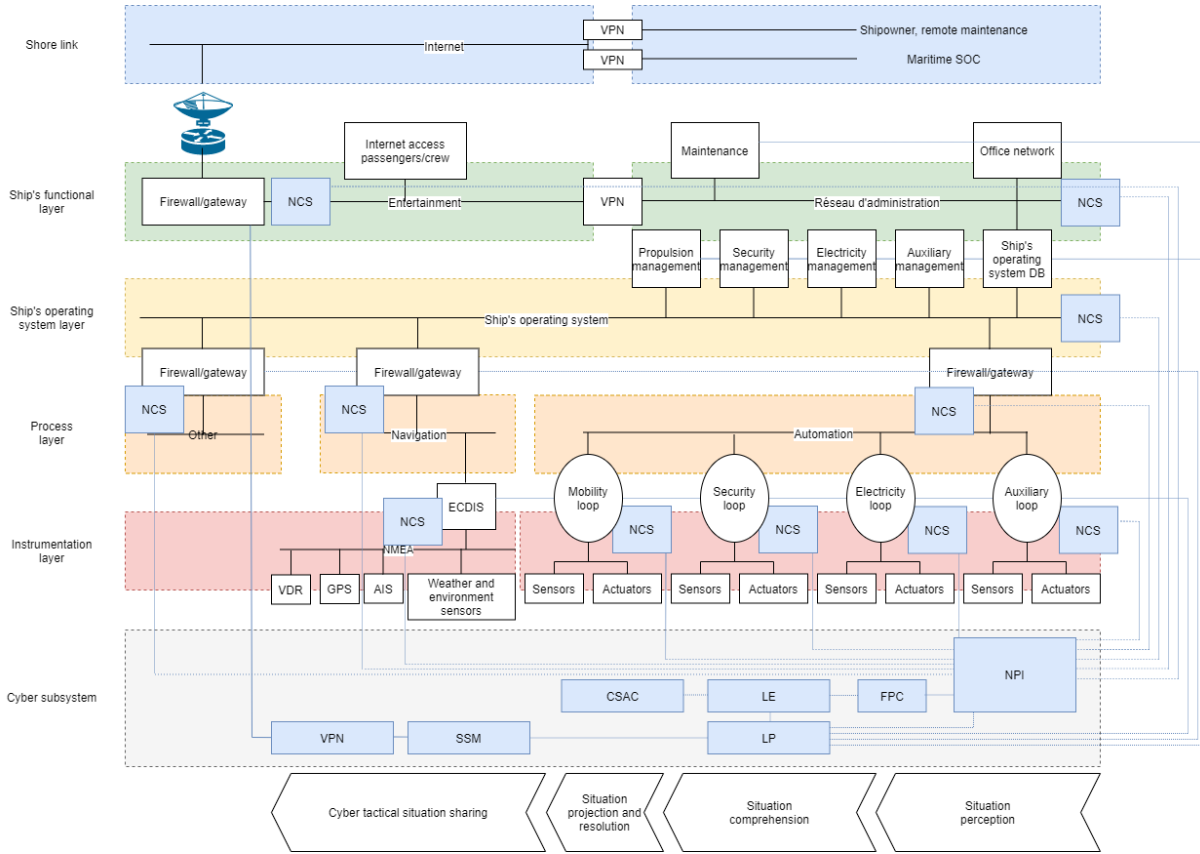


Figure IV.6: Simplified technical modeling of the interconnections of the experimentation platform with the cyber prototype subsystem (source: personal work).

concepts, we create and integrate a "cyber" subsystem into this generic ship prototype (figure IV.6). This subsystem includes the various elements necessary for the situation assessment process and for the elaboration of the MCSA, by adding NIDS and HIDS sensors, but also by retrieving logs from potential producers and adding a separate cyber monitoring layer. This work is based in particular on scientific articles published by ANSSI ([Chi14] and [Dia14]), as well as other data analysis architectures in a context of detection of cyber anomalies [Pan18]. The technical architecture has been detailed in [Jac18].

For each sub-network, a first block consisting of a sampling device like a Test Access Port (TAP) or port mirroring type is put in place. It is called Network Connexion Safety (NCS) in the article [Jac18]. This device makes it possible to avoid any "write" interaction with the monitored network.

A second device, called Network Probe Isolation (NPI), makes it possible to host several NIDSs on a single physical server by relying on the architecture and technologies proposed in [Chi14]. Each sensor is isolated from the others by the use of secure containers, which

allows separate operation of sensors using detection engines, pre-processors and separate rules. It also provides the ability to configure the sensor to retrieve metadata and metrics specific to the monitored system. Within this device, a Full Packet Capture (FPC) tool can be set up to allow for a very fine analysis directly at the packet level, which can be useful in case of in-depth investigation on particular protocols, the traffic being stored and indexed in the big data tool consisting of the Local Engine (LE).

The third block, Local Preprocessor (LP), provides input normalization, filtering and deletion of unnecessary data, as well as data transformation if necessary (for example, in the form of data enrichment with the IP address resolution). It also allows a correlation of events and alarms in order to limit the impact on the satellite link.

The fourth block, LE, allows short and long term storage of information on board the ship, thus offering an indexing and search capacity in real time on the one hand, and long retention for legal reasons on the other hand. The storage of these data and their indexing also make it possible to carry out machine learning operations.

The fifth block is the Ship Shore Manager (SSM): this tool works like a cache to guarantee that the data is correctly sent ashore even in the event of a satellite cut or insufficient speed. It limits the bandwidth allocated to the cyber subsystem to prevent it from occupying a too large percentage of the total bandwidth. It also makes it possible to prioritize the data emitted according to a given marking (priority emission of alerts over metadata, for example). If the satellite link is cut, the data is stored in cache, in memory or on a magnetic medium or flash memory, then re-transmitted to shore in the correct order when the link is reestablished.

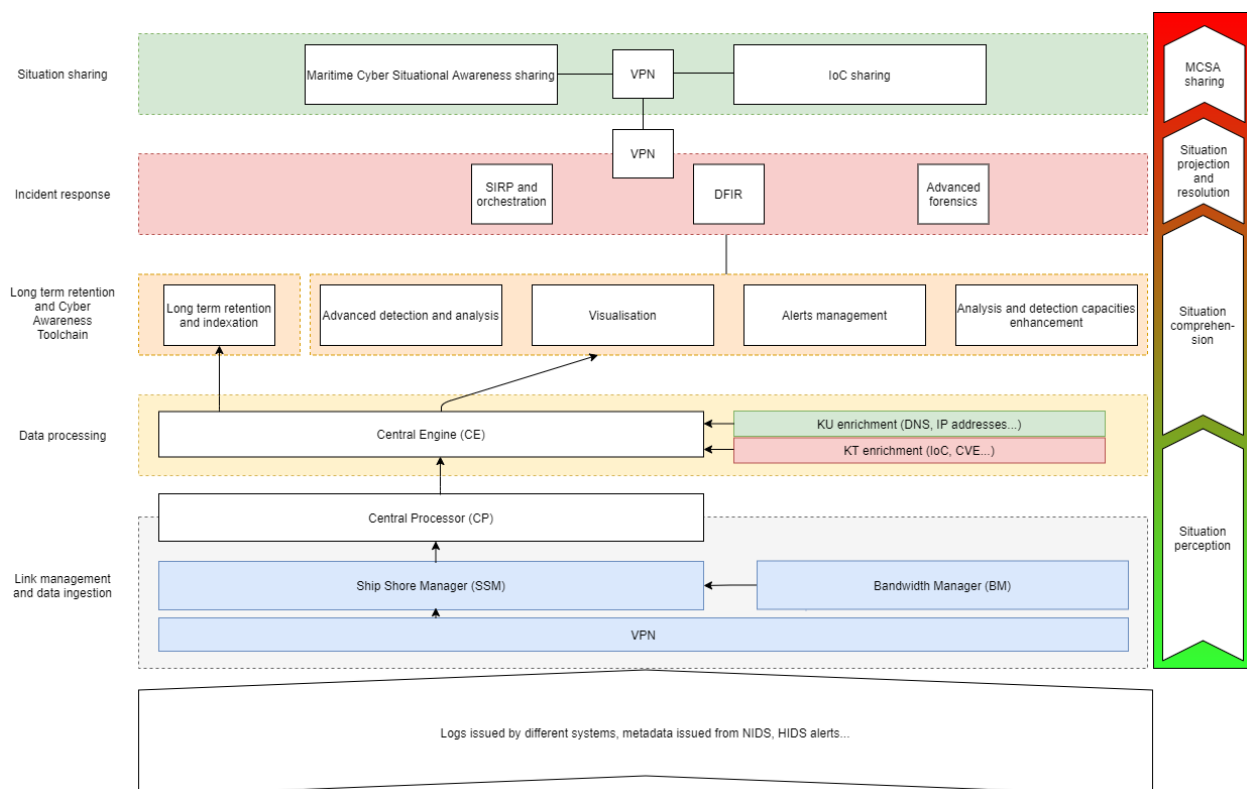
The sixth block consists of the Cyber Situational Awareness Console (CSAC). Using this console, the ship's crew has a simplified, yet complete, view of the on-board cyber situation which is shared with the SOC on shore. Even in the event of an interruption of the satellite link, it has feedbacks enabling the first diagnostics to be carried out if necessary. Finally, the link between the ship and the SOC must be secure, as it transmits critical information and must not be corrupted or captured. This is the reason why a dedicated VPN allows information to be transmitted to earth and isolates the cyber monitoring flow from other data flows.

IV.1.2 "Shore" Component

As shown in figure IV.7, on the shore side, the maritime SOC is organised around four blocks. The first block, in charge of managing the link with the ship and ingesting data,

terminates the sea/shore link with the shore SSM. It guarantees that the data collected on board is received ashore, in an orderly manner, in compliance with the fixed bandwidth constraints, even in the event of a cut in the satellite link through the Bandwidth Manager (BM) and part of the Central Processor (CP).

The third block concerns long-term retention and visualisation, with tools allowing the generation of dynamic dashboards and detailed data analysis.



This software architecture requires the implementation of data flows and capture,

analysis and presentation hardware and software, detailed in the following section.

IV.1.3 Software architecture for MCSA elaboration on the Naval Cyber Range

This subsection details the software architecture implemented for the capture of "network" and "system" data from MISs, their centralization and their analysis with the perspective to elaborate the MCSA.

IV.1.3.1 Software architecture for network data

The figure IV.8 models the end-to-end analysis of a network flow, including the part of extracting metadata and alerts, updating and improving signatures and a feedback loop information to the CSAC and the MIS.

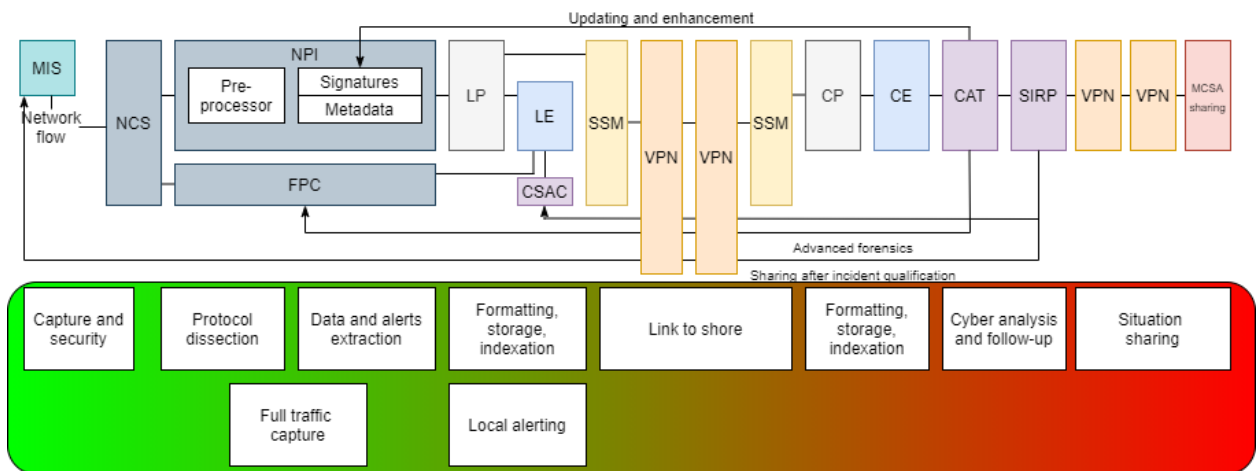


Figure IV.8: End-to-end analysis of a network flow (source: personal work).

IV.1.3.2 Software architecture for system data (logs)

The figure IV.9 models the whole analysis of a log flow resulting from HIDS or other producer, including the part extracting metadata and alerts, updating and improved signatures and a feedback loop to CSAC and MIS.

The two flows are carried out in real time and can complement and correlate each other

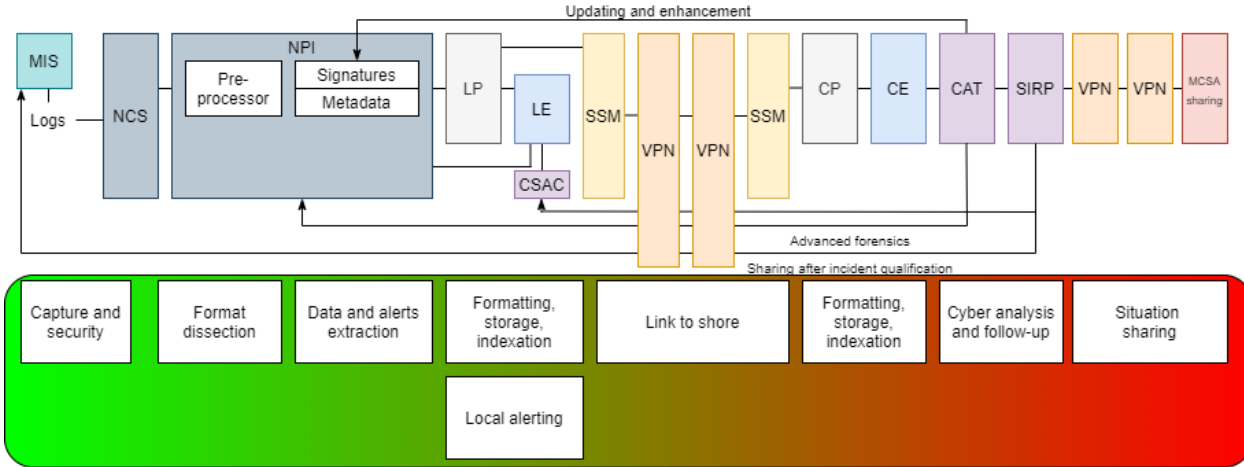


Figure IV.9: End to end analysis of the logs flow (source: personal work).

within the framework of the overall analysis of a cyberattack as modeled in Appendices C.1 and C.2.

IV.2 Quality analysis - results

Quality is one of the processes of the situation assessment (Quality Assessment) to be assessed as part of the development of the SA (see paragraph III.3.4.6). Indeed, the data analyzed by the systems in order to give them meaning and knowledge for a decision-making end (intelligence) (figure IV.10), can present quality defects (erroneous, incomplete, imprecise and uncertain data). These defects, of contextual, intrinsic or extrinsic origin [ML17], involuntary or following a tactical maneuver by the adversary, can distort the knowledge obtained of the situation and therefore impact decision-making. Measuring the quality of the data is therefore essential. In order to verify the level of quality reached on the prototype platform, quality criteria and measures were defined and then evaluated.

As summarized in figure IV.11, in our multi-sensor context in a system of systems, the objective of quality analysis is therefore to identify criteria (trustworthiness, truthfulness, completeness, freshness, timeliness), which ensure that decision-making will be based on reliable and high-quality data and information.

This section details the modalities, measurements and quality results obtained on this experimental platform developed in a context as close as possible to the reality of the ship (the satellite link is simulated realistically: latency, cuts). All the quality criteria are fed

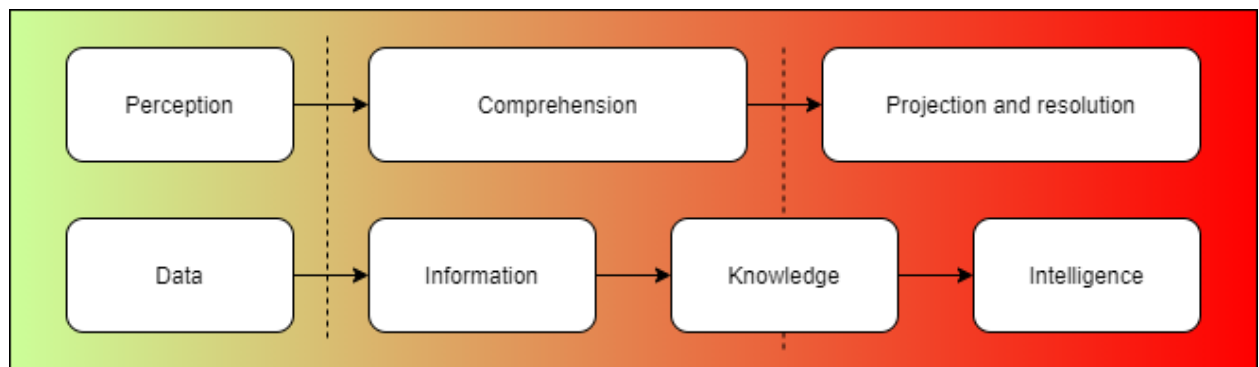


Figure IV.10: Possible parallel between data, information, knowledge, intelligence and situational awareness (source: personal work).

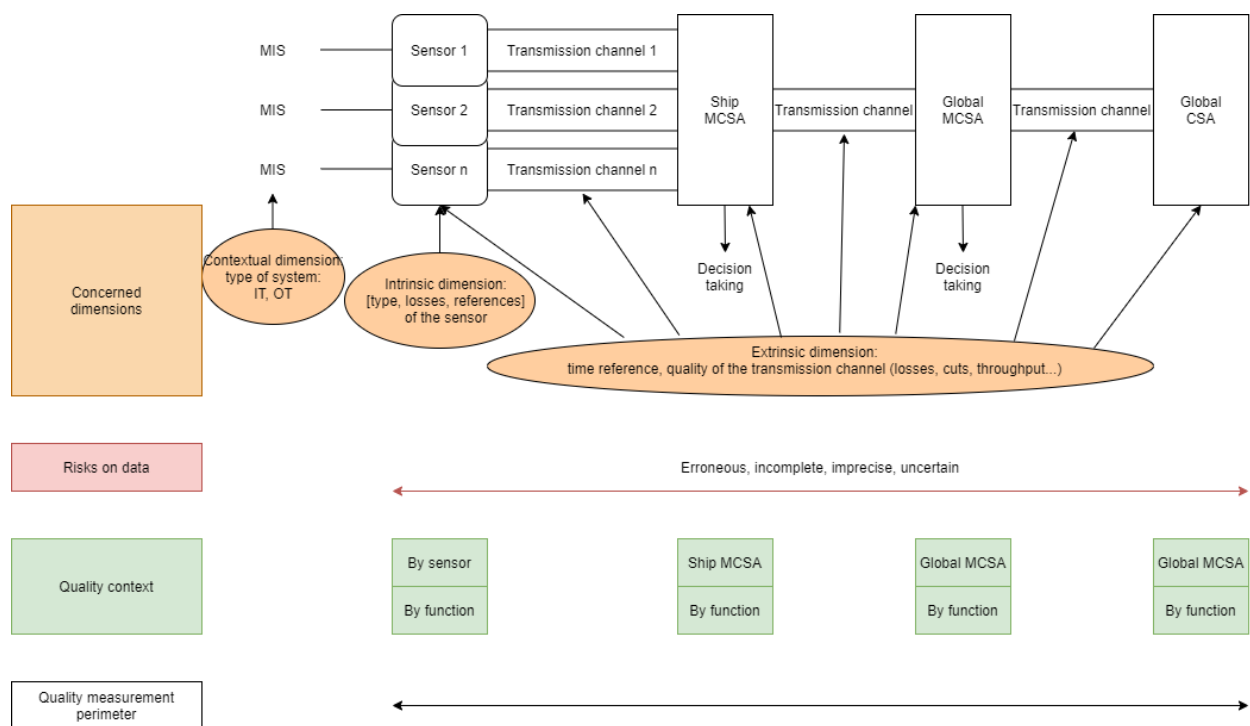


Figure IV.11: Quality measurement: overview of the dimensions of quality, links with data and risks (source: personal work).

back, in real time, such as cyber metadata, to the CP, CE and CSAC servers for information and therefore are shipped with the data throughout the process establishment of the SA. Regarding quality criteria, the detailed analysis concerns network-type sensors. However, it remains possible to identify generalities in these criteria, making them applicable to other types of cyber sensors.

IV.2.1 Quality criterion Q1 - Trustworthiness

The quality measurement linked to the trustworthiness criterion measures the confidence that one can have in the integrity of the sensor and in its ability to transmit its data in the analysis stream. This confidence is measured through four sub-criteria, measuring the effective availability of the sensor (Q1.1), its integrity (Q1.2), the timeliness of its detection signatures (Q1.3), and correct operation of the elongation link to shore (Q1.4).

IV.2.1.1 First sub-criterion of trustworthiness Q1.1: availability

First of all, the processes linked to the NIDS sensors must be functional: this capacity is measured by the availability rate of the NIDS service. This criterion is called Q1.1. With n the number of NIDS and $avail(x) \rightarrow \{0; 1\}$ the availability, Q1.1 can be expressed as a percentage in the form:

$$Q1.1 = \frac{\sum_0^n avail(i)}{n} * 100$$

It should be noted that this availability can be calculated for each level of abstraction specified in §III.3.3.

On our experimental platform, we have developed a bash script on the NIDS allowing to inform the SIEM of the proper working of the sensor (with, for example, a line like *Feb 25 09:57:02 root: NIDS on*).

IV.2.1.2 Second sub-criterion of trustworthiness Q1.2: integrity

Then, the sensors must not have been tampered with. This criterion is called Q1.2. With n the number of NIDS and $integ(x) \rightarrow \{0; 1\}$ integrity, Q1.2 can be expressed as a percentage in the form:

$$Q1.2 = \frac{\sum_0^n integ(i)}{n} * 100$$

On our experimental platform, we evaluate Q1.2 by checking, through the HIDS or through a bash script, the integrity of the configuration and data directories for our sensors. Thus, the integrity of the signatures is ensured by the regular realization of a crypto-

graphic hash, which goes back to SIEM, for verification. The line *Sep 15 16:01:51 nids-mob pi: SHA256 bd819ce1cb2ddfa3ddad588971da099c2e873b4a143a3cadd61c3b685e0985c2/etc/suricata/rules/smtp-events.rules* thus returns the crypto hash of the Simple Mail Transfer Protocol (SMTP)-related signatures of the NIDS sensor, thus guaranteeing its integrity by comparison with a reference list of cryptographic hash of signatures.

IV.2.1.3 Third sub-criterion of trustworthiness Q1.3: signatures freshness

Then the signatures must be up to date. This criterion is called Q1.3. With n the number of NIDS and $freshness(x) \rightarrow \{0; 1\}$ the freshness of signatures Q1.3 can be expressed as a percentage in the form:

$$Q1.3 = \frac{\sum_0^n freshness(i)}{n} * 100$$

On our experimental platform, Q1.3 is measured by two parameters obtained and transmitted on a regular basis: the first consists of the date of the signature files (for example: *Feb 24 16:01:51 nids-mob pi: Date /etc/suricata/rules/smtp-events.rules: Friday, December 21, 2018, 15:43:33 (UTC + 0100)* indicates that the creation date of the file */etc/suricata/rules/smtp-events.rules* is December 21, 2018 at 15:43:33) and the second concerns checking some parameters of the NIDS: *stats.detect.engines.last_reload*, which returns the last rule reload date, *stats.detect.engines.rules_loaded*, which specifies the number of rules loaded, and *stats.detect.engines.rules_failed*, which gives the number of rules in error. Thus, a reload date of the rules prior to the date of the signature files will inform that the last version has not been loaded, for example by means of an alert.

IV.2.1.4 Fourth sub-criterion of Q1.4: shore-link availability

The availability of the shore link is essential to ensure that the link with the SIEM is well assured. This criterion is called Q1.4. Formally, with *MTBF* the Mean Time Between Failure and *MTTF* the Mean Time To Failure, the availability D of a system is usually written as:

$$D = \frac{MTTF}{MTBF}$$

The availability measure can be obtained by simulating a complete connection with the SIEM (test of the service), or by means of a request using the Internet Control Message Protocol (ICMP) protocol (test of the network). As the need emanates essentially from the detection of a satellite cut, we have retained the second possibility. By creating a bash script, we get in the logs a new recurring syslog entry, initiated for example every minute: "*Feb 24 14:29:01 nids-elec root: MCSA server accessible*". In addition to the other indicators, this measurement allows us to identify any possible break in the link.

IV.2.2 Quality criterion Q2 - *Truthfulness*

The second quality criterion, Q2, aims to verify the veracity and solidity of the input elements (truthfulness or soundness). In certain cases, in particular in the case of detection, the mechanisms used are generally based on signatures. These signatures, if they prove to be incomplete, can lead to false positives and, if they are incomplete, to false negatives (*cf* §III.3.4.1). The case of false positives is symptomatic of the daily difficulties of SOC's [KSB⁺19]. These false positives, which may be due to poor quality signatures, can also be intentionally generated by an attacker for deception. Indeed, faced with a storm of false positives intended to distract the defender's attention from the real target, the detection and prioritization of a real cyberattack is proven to be complex.

The quality of this criterion rests on several bases. First, the NIDS sensor must be able to analyze the protocol used to carry out the cyberattack. Then, the signatures used by the sensor (generic, like *Emerging Threats*, specific, for example by the reception of IoC sent by a CSIRT, or descriptive, describing the nominal network traffic present on the network), however precise they may be, do not necessarily allow the detection of the exploitation of a zero day type flaw. Finally, cyberattack network datasets are still relatively few in number, sometimes incomplete or obsolete, and often poorly suited to meet our maritime and industrial context [HBB⁺20]. It therefore does not appear relevant to generate an dedicated indicator for this criterion Q2: the signatures made available to the sensors must be considered as "at the state of the art" in a context of detection by signature, and this verification is already carried out for criterion Q1.3.

IV.2.3 Quality criterion Q3 - *Completeness*

The third quality criterion, Q3, must verify the completeness of the input elements (completeness). Indeed, the sensors, in particular the network, may ignore or not be able to reconstruct certain input packets which may, for example, hamper their ability to subsequently reconstruct complete files. The link to shore and satellite bandwidth constraints can also be factors in data loss. The objective is therefore to measure Q3.1, corresponding to the loss rate of the NIDS sensors, Q3.2, the loss rate due to the satellite link and Q3.3, indicator which verifies the proper working of the capture.

IV.2.3.1 First sub-criterion of completeness Q3.1

There are several reasons why a NIDS might fail to analyze all of the packets passing through the network it is listening to. They can be due to the very high speed of the local network, causing limitations in terms of memory or overloading of processor resources, cyclic redundancy checksum errors, etc. With nb_{rp} the number of packets received by the NIDS and nb_{ip} the number of invalid packets, we can formalize Q3.1 using precision in the form:

$$Q3.1 = \frac{nb_{rp} - nb_{ip}}{nb_{rp}} * 100$$

Packet loss can be measured at the network card itself, for example with the use of the *ethtool* command on a *GNU/Linux* system. We assume that this measurement was carried out when the sensors were installed to detect any hardware anomaly. To analyze the possible loss of packets at the level of the NIDS itself, we track the value of the parameter *suricata.eve.stats.capture.kernel_drops* over a period of one week, in order to deduce a percentage loss compared to the number of captured packets (*stats.capture.kernel_packets*). During the week of monitoring, the value of this parameter on all the sensors of our experimentation platform, we did not observe any change in this figure during our measurements: it thus remained constantly at 0, the indicator of *completeness* Q3.1 thus always being equal to 100%. However, it appears relevant to continue to measure this parameter over time, in order to detect any unfavorable development.

IV.2.3.2 Second sub-criterion of completeness Q3.2

The speeds of satellite links can vary significantly between ships, depending on their type, their needs and their geographical position. The vast majority of them are asynchronous, the reception requirements on board the ship generally being greater than the transmission requirements. The maximum bit rate available at a t time is essentially linked to questions of subscription cost, technology, the diameter of the antenna on board the ship, the number of antennas and environmental factors. As technology has made significant progress in recent years, downlink speeds range from 8 Mbps for a cruise ship to 6 Mbps for a ferry, 2 to 4 Mbps for a merchant ship, and 2 Mbps/s for a modern fishing vessel, sometimes less. Recent tests give hope for a strong increase in bandwidth and speed for the years to come². The upstream speeds are generally related by a factor of 10 to the downstream speed (for example: 256 kbits/s upstream for a downstream speed of 2 Mbits/s³). Bandwidth for the elaboration of the MCSA are essentially rising, in order to ensure the proper upload of relevant metadata, while the downstream speeds are lower and essentially linked to the downloading of new signatures or to the sharing of information.

Satellite links are constrained, whether by bandwidth, transmission and reception delays, errors which are generally the consequence of difficult environmental conditions (rain, for example) or of link congestion. On our experimental platform, the simulated satellite connection has, by default, the characteristics of a fixed link. The figure IV.12 shows an average delay of less than 1 ms which is not realistic at all for our target maritime environment.

The theoretical average Round Trip Time (RTT) of an end-to-end link for a satellite in geostationary orbit (which is the case for most satellites used in the maritime world, such as satellites used for Very Small Aperture Terminal (VSAT) services) is of the order of 500 ms. Indeed, the distance to be traveled by radio waves between the earth and a geostationary satellite is approximately 36,000 km. At a speed of light of 299,792 km/s, the time between transmission on land and reception at the satellite level is therefore 0.12 s. For a earth/satellite/earth round, the time is therefore about 250 ms and 500 ms for the complete RTT between the transmitter and the receiver. If we add to this value the time necessary to reach the teleport router and the distribution on board the ship (figure IV.13), the RTT is therefore greater than this optimal value and, from the measurements we were able to do, generally closer to 650ms to 800ms under normal conditions.

²<https://www.satellitetoday.com/mobility/2018/02/26/watch-ses-ceo-setting-new-bandwidth-capacity-record/>

³<https://www.nsslglobal.com/index.php?idPage=4&p=101433>

```

ping 11.0.0.2
Sending 'Ping' request to 11.0.0.2 with 32 bytes of data:
Answer from 11.0.0.2: bytes=32 delay<1ms TTL=63
Answer from 11.0.0.2: bytes=32 delay<1ms TTL=63
Answer from 11.0.0.2: bytes=32 delay<1ms TTL=63
Answer from 11.0.0.2: bytes=32 delay<1ms TTL=63

Ping statistics for 11.0.0.2:
    Packets: sent=4, received=4, lost=0 (0% loss),
Approximative loop delay in milliseconds:
    Minimum=0ms, Maximum=0ms, Average=0ms

```

Figure IV.12: Average delay of an ICMP *ping* on the WAN of the *Naval Cyber Range* platform (source: personal work).

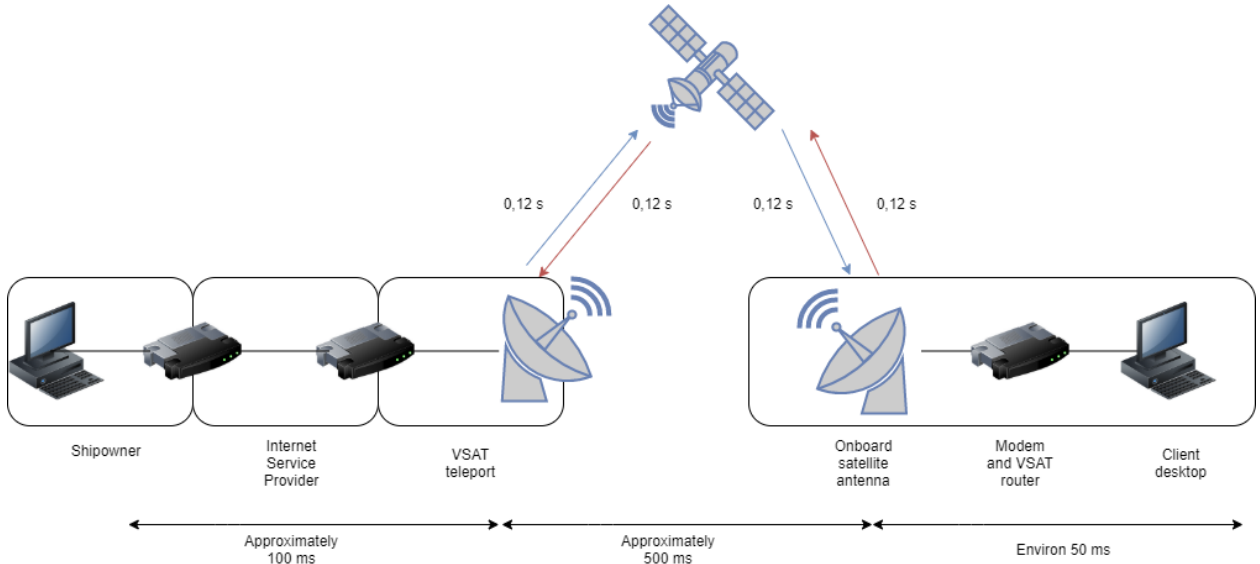


Figure IV.13: Diagram of the Round Trip Time of a satellite link to a ship at sea (source: personal work).

We therefore add a delay of 650 ms and a normal distribution of ± 20 ms to this delay. As the satellite link is not exempt from packet loss, we inject an unfavorable condition which corresponds to a loss of 10% of packets. In order to simulate these features on our platform, we use the GNU/Linux system and the *tc* command. *tc*, for *traffic control*, is a utility which allows to configure the GNU/Linux packet scheduler and, in our case, to degrade certain characteristics to simulate a satellite link. Thus, the command *tc qdisc add dev ens19 root netem delay 650ms 20ms distribution normal loss 10% corrupt 5% duplicate 1%* allows us to simulate adverse conditions on a simulated satellite link, as confirmed by the results of

figure IV.14.

```

Answer from 11.0.0.2: octets=32 delay=636ms TTL=63
Answer from 11.0.0.2: octets=32 delay=652ms TTL=63
Answer from 11.0.0.2: octets=32 delay=629ms TTL=63
Answer from 11.0.0.2: octets=32 delay=650ms TTL=63
Timeout exceeded
Answer from 11.0.0.2: octets=32 delay=626ms TTL=63
Answer from 11.0.0.2: octets=32 delay=670ms TTL=63
Answer from 11.0.0.2: octets=32 delay=641ms TTL=63
Timeout exceeded

Ping statistics for 11.0.0.2:
    Packets: sent = 37, received = 33, lost = 4 (10% loss),
Approximate loop delay in milliseconds:
    Minimum=603ms, Maximum=676ms, Average=649ms

```

Figure IV.14: Adapted ICMP *ping* delay on the WAN link of the platform (source: personal work).

The *tc* command also allows to constrain the transmission and reception bandwidth, as well as the Maximum Transmission Unit (MTU), by the use of the *rate* parameter⁴. On our prototype, we therefore constrain the downlink bandwidth to 2 Mbits/s, the upstream bandwidth to 256 kbits/s. We also adjust the MTU and the link degradation parameters in order to best approximate the characteristics of a satellite link of a medium-sized ship.

On this link, now degraded by simulated outages which generate packet loss, delays and corruption, we send our cyber monitoring flow from the ship to land to verify the reliability of the architecture and the absence of loss of information due to the satellite link. To ensure data completeness, we perform a cryptographic checksum on both incoming and outgoing data. We note that the checksums on arrival and departure are identical: by design, the defined architecture therefore guarantees the absence of data loss on the satellite link. We have also voluntarily interrupted the data link on several occasions. Despite these cuts, all of the data has been received, the SSM onboard side ensuring the data cache (within the limit of the size of its hard disk), as long as they have not been received by the SSM on the shore side.

It should also be noted that it is not conceivable to use all of the satellite bandwidth for

⁴For instance, by the command *tc qdisc add dev ens19 parent 1: handle 2: tbrate 256 kbit burst 256k latency 50ms mtu 1460* for the uplink and *tc filter add dev ens19 protocol ip parent ffff: prio 1 u32 match ip src 0.0.0.0/0 policy rate 2Mbit burst 2Mbit drop flowid: 1* for the downlink.

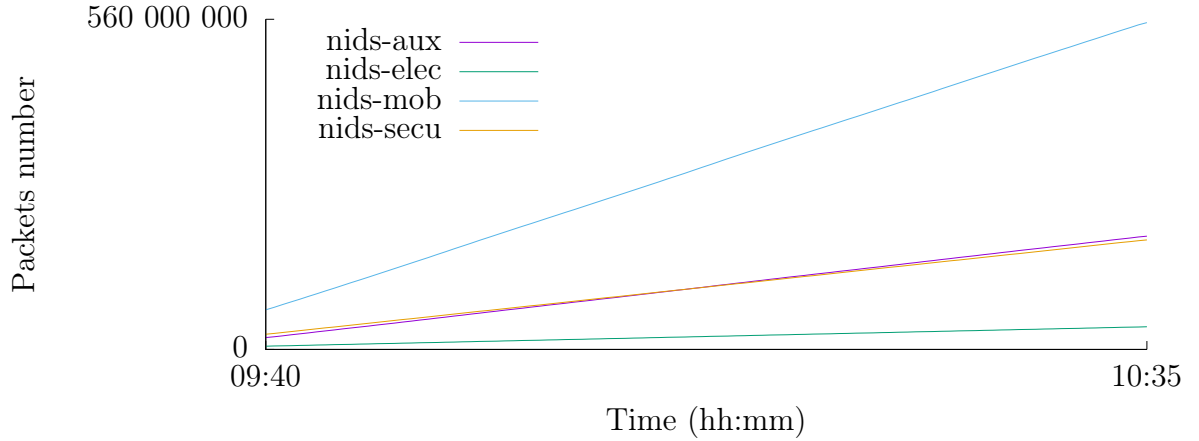


Figure IV.15: Evolution of the *stats.capture.kernel_packets* parameter for the four industrial loops (source: personal work).

the sole purpose of cyber anomaly detection. It is therefore necessary, within the satellite link, to also constrain the link between the SSM. The precise value of the necessary bandwidth depends on the type of vessel, the number of sensors and their settings, as well as the volume of data passing through the vessel. On our prototype, we were able to restrict the bandwidth up to $1/32$ of the bandwidth without this causing a noticeable inflection of the criterion.

The architecture having proven to be, by design, qualitative, we therefore propose not to follow any parameter associated with this quality criterion.

IV.2.3.3 Third sub-criterion of completeness Q3.3

On our prototype platform we studied, for each loop, the evolution over time of the *stats.capture.kernel_packets* value. This value, returned by the NIDS sensor, indicates the number of network packets captured by the GNU/Linux kernel. We follow this value for the NIDS sensor present on each loop (*nids-aux*, *nids-elec*, *nids-mob*, *nids-secu*). The objective of this measurement is to confirm the interest of monitoring the parameter for establishing indicator Q3.3, but also to verify the deterministic nature of industrial systems of the OT type.

Figure IV.15 presents the measurements of indicator Q3.3. The ordinate indicates the value of the parameter *stats.capture.kernel_packets* and the abscissa the time, the sampling being fixed at 1s.

It's interesting to notice that the relation between the *stats.capture.kernel_packets* pa-

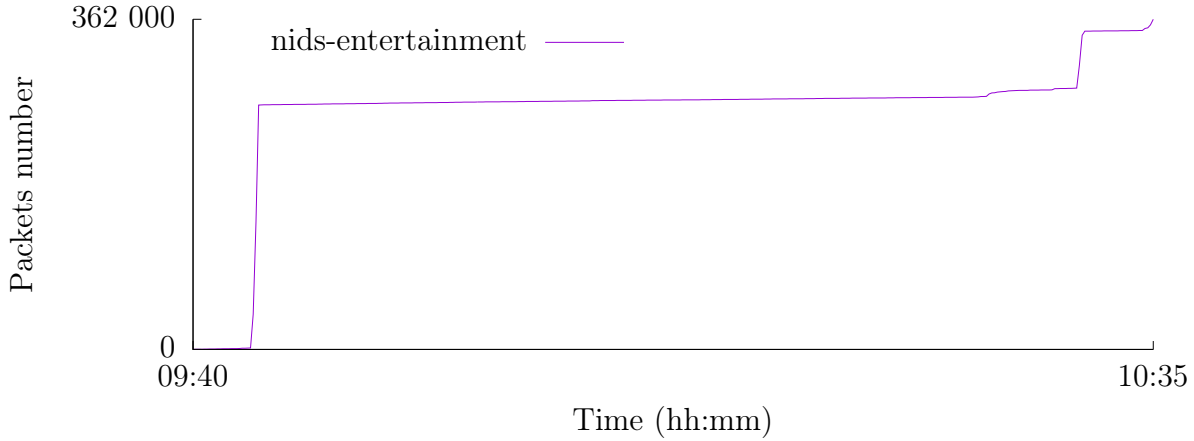


Figure IV.16: Evolution of the *stats.capture.kernel_packets* parameter of an IT network (source: personal work).

Table IV.1: α directing coefficient of the half line representing the *stats.capture.kernel_packets* parameter on the four industrial loops

Sensor	Directing coefficient α	PLC	I/O Nr	Variables
Electricity loop NIDS	10,32	4	39	268
Security loop NIDS	45,02	4	49	296
Auxiliary loop NIDS	54,59	4	50	302
Mobility loop NIDS	160,43	3	63	540

parameter and time approaches a linear function $f(x) = \alpha x$. The explanation is linked in particular to the relatively deterministic nature of so-called "industrial" networks. Indeed, the OT is more characterized by the functioning of the ship itself and relatively little influenced by human activities. *A contrario*, as we can see on a IT system (*cf* figure IV.16), the same parameter evolves in a non-deterministic way, or in any case on a determinism based on the user activity on the network (downloading files, videos, limited activity at night, etc.). This evolution therefore reflects both the latent and permanent activity of the information system, but also, mainly in the case of IT, human behavior impact on network activity.

Table IV.1 indicates the directing coefficient α of the function f representative of the variation of the parameter *stats.capture.kernel_packets*.

The director coefficient α seems to be linked to the number of analog and digital inputs and outputs per loop, but above all to the number of read and write operations of variables per loop.

To design indicator Q3.3, we worked on time windows to check that, on the time series,

the value is stable between two windows. The measurement was carried out on a sliding time window of 30 seconds. The measurement of the directing coefficient α at t and $t + 30s$ is done on the 4 loops. By simulating a loss of capture on a loop following a hardware failure, the f function is no longer linear but becomes constant, gradually or suddenly: the directing coefficient α tends towards 0 or becomes zero: the degradation is then immediately visible. Figure IV.17 shows the results of this test. The degradation of Q3.3 appears on the right of the graph: α then tends towards 0. The point of vigilance remains the width of the time window: too wide, it risks not showing short capture faults. Too short, it risks raising too many false positives. In our work on the platform, the value of 30 seconds gave the best results.

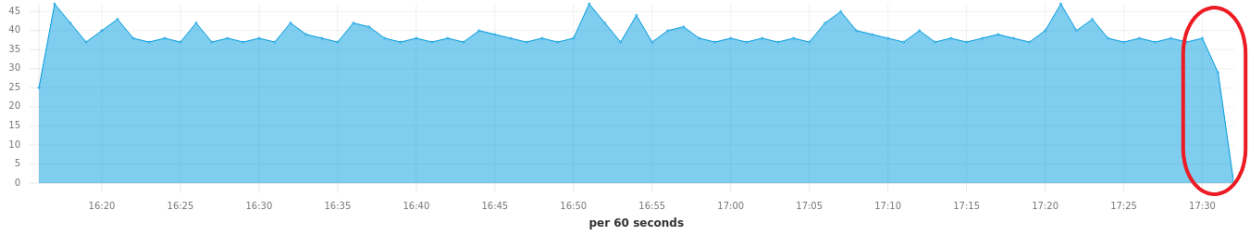


Figure IV.17: Deterioration of the sub-criterion Q3.3 following a hardware failure simulation (source: personal work).

IV.2.4 Quality criterion Q4 - *Freshness* and *timeliness*

The fourth quality criterion, Q4, checks the freshness of the input data. The good ordering of the input data (timeliness) is also essential so that they are presented for analysis in the right order, in order to allow the creation of chronologies during the phases of digital investigation.

For the freshness criterion, we establish an indicator corresponding to the difference between the timestamp on insertion into the data lake (LE) ($t_{@timestamp}$) and the capture timestamp ($t_{timestamp}$) by the NIDS:

$$Q4 = t_{@timestamp} - t_{timestamp}$$

Q4 is calculated, in the LE, by the *Painless* formula given in figure IV.18.

Freshness is essentially dependent on the proper functioning of the satellite link (no cut-

```
if(!doc['timestamp'].empty) {
  return (doc['@timestamp'].value.getMillis() -
    doc['timestamp'].value.getMillis())/1000 }
}
```

Figure IV.18: Calculation formula of the freshness criterion Q4 (source: personal work).

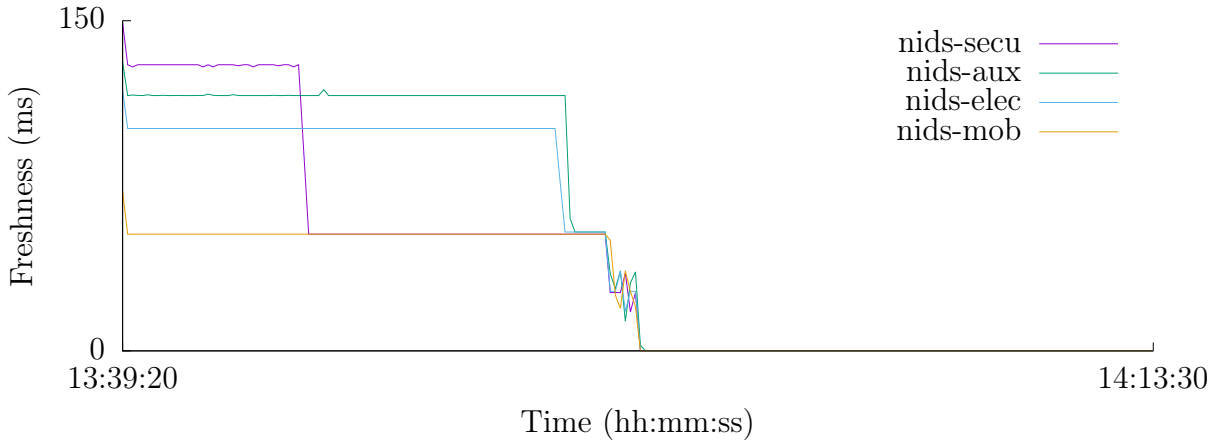


Figure IV.19: Positive evolution of the data freshness criterion during the NTP synchronization phase (source: personal work).

off) and the available bandwidth. Thus, a metadata captured at a time t , but which could not have been transmitted immediately will see its freshness criterion less good than that of a metadata returned immediately between the MIS and the SIEM on land, which will have better freshness (around 350 ms, depending on the delay of the satellite link). Accuracy obviously requires perfect synchronization of the clocks of the sensor and the SIEM, for example by using a reliable time reference such as GNSS.

As part of our prototype, we produced a real-time graph which allows us to follow the evolution of the freshness quality parameter on board the simulated ship and at the shipowner's SIEM level. Figure IV.19 shows the Network Time Protocol (NTP) synchronization phase. Initially, the sensors and the SIEM are not synchronized from a time point of view (from the start of the graph until approximately 13h55min00s). The NTP synchronization gradually takes its effect to arrive, starting at about 13h56min00s, at an almost perfect synchronization and at an excellent level of freshness. The measurement of the freshness would make it possible to quickly establish, beyond a clear failure and all other things being equal, a lower quality of the clock of a sensor.

The purpose of the timeliness quality measurement is to check that, even in the event of a satellite outage, the data received after the outage are indeed exhaustive and that they are

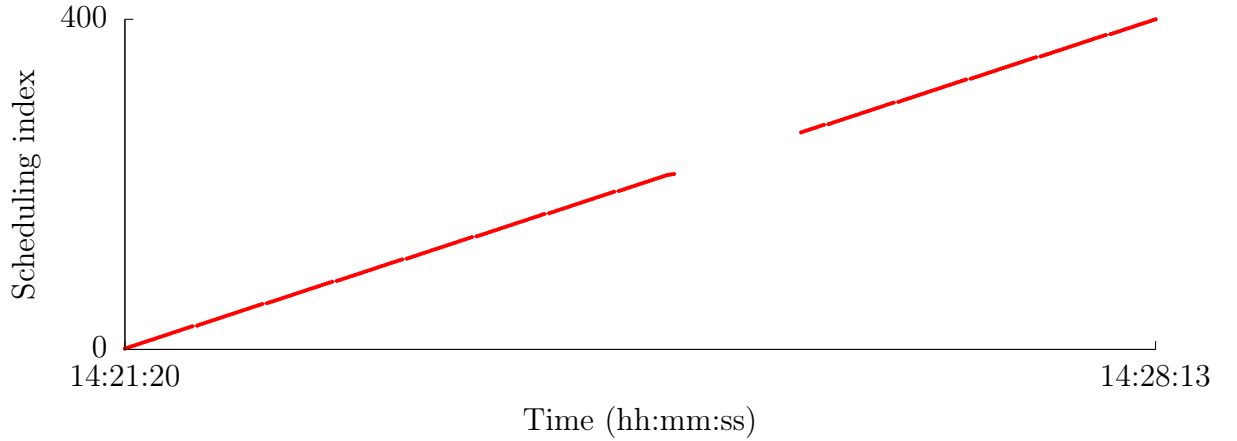


Figure IV.20: Data timeliness following a satellite outage simulation (source: personal work).

reindexed with the correct timestamp. If this were not the case, they could end up missing or scattered at timestamps that do not match their generation timestamp (for example: their indexing time in the CE). To perform this measurement, a file comprising a million events is generated. Each event is separated by one second, and ordered by an index incremented at each event. We then send this file to the on-board local engine and to the shore local engine, which ends the satellite link. On arrival, we check that the message is exhaustive, that the data has arrived in order and that it is correctly reindexed. The figure IV.20 makes it possible to identify the regular progress of the reception of the index at the LE level until the satellite outage (14h25min00s, time of the SIEM). The graph seems to indicate a loss of the corresponding indexes. However, the analysis of the received data (table IV.2) shows that the corresponding data is well received, the only difference being that some $n+1$ indexes were received at the same time as n indices. The possible misinterpretation is linked to the fact that the time stamping of the tool is carried out on the time reference of the SIEM and not of the sensor, on the one hand, and to the graph tool which does not represent correctly the exact reception of the data. By using the time reference of SIEM and a shorter time window, we then check that the corresponding data has been received (see, figure IV.21, the recovery of the delay, circled red). The use of the sensor timestamp therefore makes it possible to fix this situation: the data transmitted is indexed to the value of the sensor timestamp and not to that of the SIEM, guaranteeing a consistent chronology: the criterion of *timeliness* is thus guaranteed by design.

Table IV.2: Data timeliness following a satellite outage simulation (source: personal work).

Timestamp	Timeliness
March 4, 2020 14:25:14.661	231
March 4, 2020 14:25:13.659	230
March 4, 2020 14:25:12.656	229
March 4, 2020 14:25:12.656	227
March 4, 2020 14:25:12.656	228
March 4, 2020 14:25:12.655	225
March 4, 2020 14:25:12.655	226
March 4, 2020 14:25:12.654	223
March 4, 2020 14:25:12.654	224
March 4, 2020 14:25:12.653	222
March 4, 2020 14:25:12.641	221
March 4, 2020 14:25:12.640	219
March 4, 2020 14:25:12.640	220
March 4, 2020 14:25:12.639	217
March 4, 2020 14:25:12.639	216
March 4, 2020 14:25:12.639	218
March 4, 2020 14:25:12.638	215

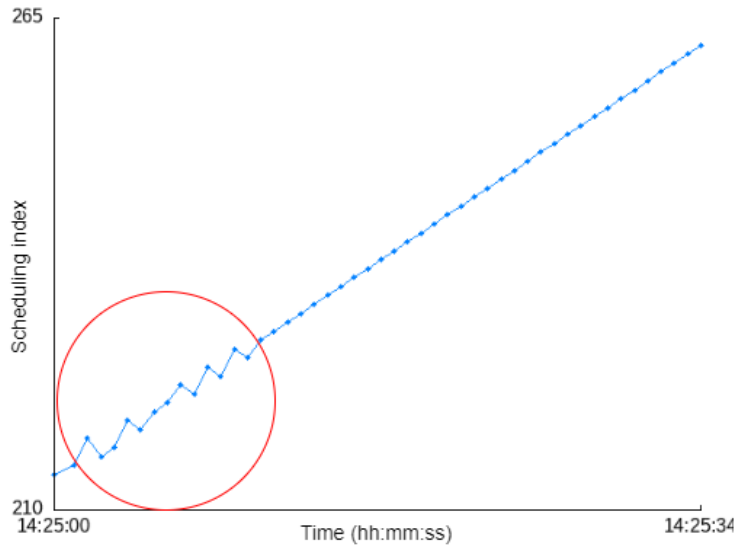


Figure IV.21: Recovery of the delay in the counter reception (source: personal work).

IV.2.5 Quality analysis - Conclusion

The work carried out allowed, beyond the realization and the instrumentation of a realistic maritime cyber range for prototype purposes, to bring added value and important

details on the notions of quality applied to the context of cyber anomaly detection on physical maritime and cyber systems. When appropriate, behind each generic term (trustworthiness, truthfulness, completeness, freshness, timeliness), we have proposed, formalized, generated and measured these values in real time on our platform, in order to have, at all times, a quality measurement essential for the development of the MCSA and for decision-making. We have seen that, in the context of the MCSA, the quality of the criteria of completeness, freshness and timeliness are essential compared to a more generic CSA outside the maritime context. We have demonstrated that our architecture makes it possible to guarantee this quality or, failing that, to measure its deviations.

Taking into account this quality and the specific features of the monitored systems (IT/OT) eases the use and analysis of data issued from cyber tools of varying design, origin and goals. The confidence, accuracy, freshness of the data they produce can be tracked to determine the quality of information needed for decision making. This prototype phase, essential for the validation of our model, now allows us to guarantee the quality of the data we collect and to continue our work.

IV.3 Visualisation and mapping - results

In this section, we answer the research question RQ2 on its visualisation aspects, and part of the RQ3 on anomaly detection on MIS as well as the points mentioned in the section III.4. In order to validate the contributions of the MCSA, all of our work will be carried out on our prototype platform.

IV.3.0.1 Dynamic passive mapping

The information that can be passively collected on our cyber range is particularly varied. Some may be common to all IT and OT systems (Layer 2 physical addresses, some Layer 3 metadata). Others are specific to industrial systems, but can be found on shore-based systems (S7, Modbus or other protocols). Finally, some are specific to MIS (NMEA standards, etc.). The real-time analysis of the network, as can be done with the proposed architecture, also makes it possible to obtain information on the software installed, by the interception of banners, negotiations in a completely passive manner. This type of technology, often called fingerprinting, is of major interest in the context of maritime systems, in

order to significantly improve the KU of MIS.

In this context, the difficulty is not so much the collection of data (situation perception) as their analysis (situation comprehension). First, a dynamic set of graphic representations was designed. It makes it possible to merge, enrich and display the relevant information relating to a piece of equipment (for example, a machine or a PLC), for mapping and anomaly detection purposes, from a source multi-sensor data. The figure IV.22 thus shows the type of visualisation which can be carried out and which represents a significant gain in terms of time for the analyst who then obtains an efficient and immediate representation of the data at his disposal. In a single real-time table, all of the passively captured information is thus presented. The information available includes, for example, the distribution of the network protocols used by the machine, the *HTTP User-Agent*, the alerts detected by the sensors and relating to the machine and their distribution over time, the source/destination ports used and the volume of traffic concerned, the DNS requests sent by the machine and the number of associated requests, or the machines with which it exchanges the most, in transmission and in reception.

IV.3.0.2 Network flow visualisation

Parallel coordinate graphs are particularly suitable for visualizing network data flows [CTSdG12]. Indeed, from a cognitive perspective, a network connection between a device A and a device B is easily represented by a segment. If this type of graph is generally carried out *a posteriori* on collected data, our architecture allows it to be carried out, in real time, on remote MIS, i.e. on data from layers 2, 3 and 4 of the Open Systems Interconnection (OSI) model, or on the application layers, if the data can be properly collected and interpreted by the sensors used.

In order to generate those graphs, we use the architecture described in figure IV.6 and carry out requests at the CE level of the SOC. Given the number of sensors positioned, we select a specific sub-network each time, in order to be able to represent the exchanges in real time. The data obtained at the flow level make it possible to correctly separate the direction of the connections (client/server and server/client), so as to correctly produce the graph, as can be seen in the figure IV.23. However, it should be noted that the activation of the netflow or flow data in the on-board sensors represents a significant overhead for the upstream satellite link, which impacts the other flows of the ship which use this same link. But it is, at present, the only way to have the information of the direction of the connection,

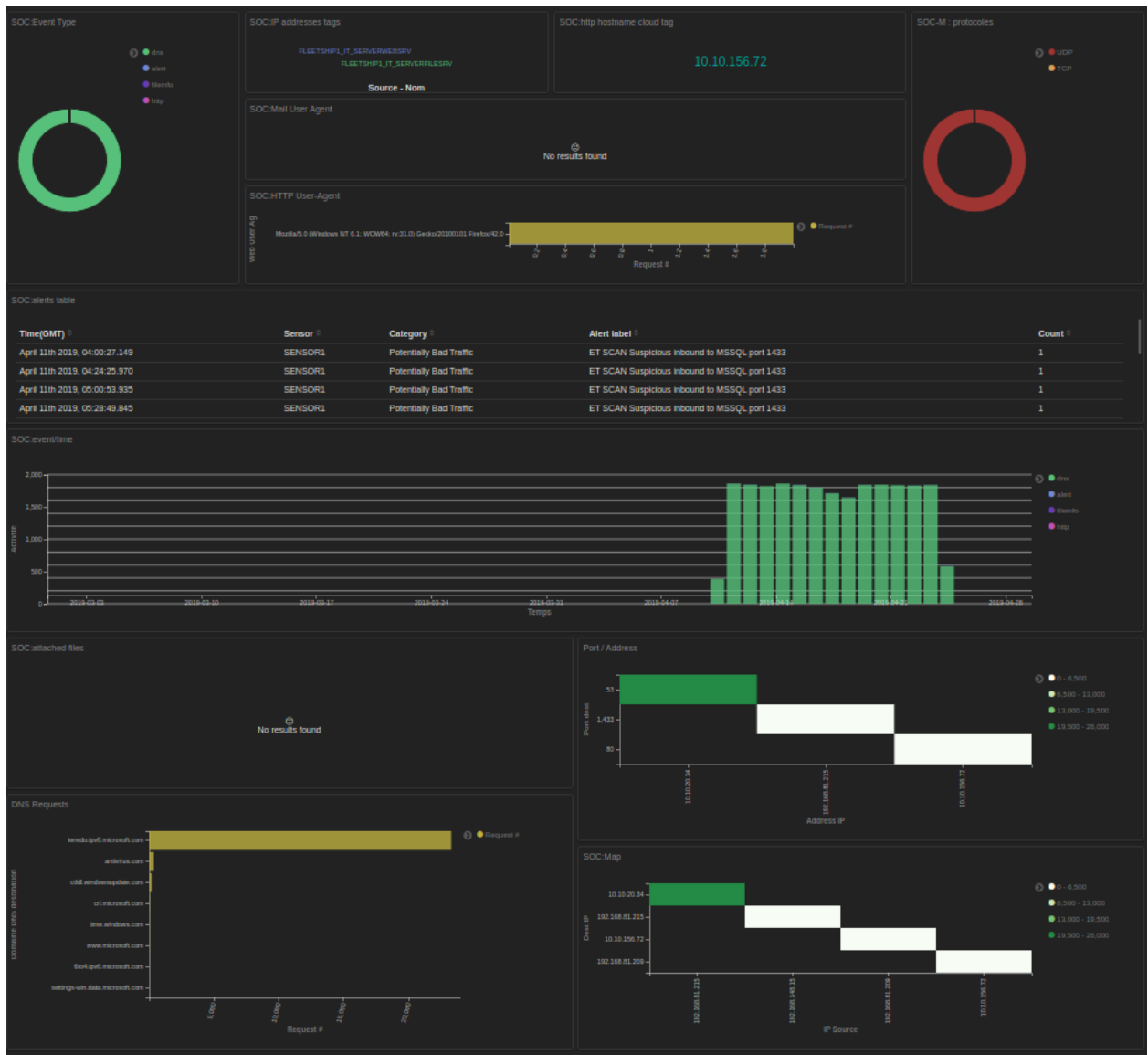


Figure IV.22: Multisensors data fusion (NIDS, HIDS) for anomaly detection on a MIS (source: personal work).

this information being absent in the application pre-processors of NIDS (as can be seen in figure IV.24).

The consequence of this absence is that the destination port is not directly exploitable to produce, for example, a parallel coordinate graph. Indeed, the distribution of destination ports (*dest_port*) mixes source ports and destination ports, as seen in figure IV.25, in the form of scatter plot: the destination ports from the answer from the server to the client are grouped together at the top (minimum ports > 1024), whereas they are in fact source ports following the three-way handshake of the Transport Control Protocol (TCP). Likewise, the

```
dest_ip: 34.95.71.207
src_ip: 192.168.1.44
@timestamp: Aug 22, 2019 @ 13:14:27.281
app_proto: tls
timestamp: Aug 22, 2019 @ 13:14:27.069
dest_port: 443
tcp.psh: true
tcp.ack: true
tcp.state: established
tcp.tcp_flags_tc: 1a
tcp.syn: true
tcp.tcp_flags_ts: 1a
tcp.tcp_flags: 1a
proto: TCP
@version: 1
tags: suricata
event_type: flow
src_port: 33,496
flow_id: 654,490,614,960,586
flow.state: established
flow.end: Aug 22, 2019 @ 13:13:28.690
flow.pkts_toserver: 24
flow.pkts_toclient: 21
flow.age: 59
flow.bytes_toclient: 2,668
flow.start: Aug 22, 2019 @13:12:29.987
flow.bytes_toserver: 3,171
flow.reason: shutdown
flow.alerted: false
```

Figure IV.23: Overview of the collected data with the flow capacity activated on a NIDS (source: personal work)

actual destination ports of the servers are visible at the bottom of the graph (*dest_port* usually < 1024).

As can be seen in the figure IV.26, this defect results in obtaining a graph with parallel coordinates having a "mirror" effect at the level of the source and destination ports, and a mixture of the source and destination IP addresses. We propose a solution which propagates the meaning of the connection at the application level datagrams (layers ≥ 4 of the OSI model) [Jac19a]. Its integration into the NIDS Suricata is in progress⁵. It will

⁵<https://github.com/OISF/suricata/pull/3521>

```
in_iface: wlp2s0
dest_ip: 192.168.1.44
src_ip: 192.168.1.254
@timestamp: Aug 22, 2019 @ 13:12:30.274
timestamp: Aug 22, 2019 @ 13:12:29.987
dest_port: 36,211
proto: UDP
@version: 1
tags: suricata
event_type: dns
src_port: 53
flow_id: 1,815,519,059,369,073
dns.rd: true
dns.grouped.A: 34.95.71.207
dns.grouped.CNAME: prod.webextstoragesync.prod.cloudops.mozgcp.net
dns.version: 2
dns.id: 6,680
dns.answers: {
  "rrname": "webextensions.settings.services.mozilla.com",
  "rrtype": "CNAME",
  "ttl": 85,
  "rdata": "prod.webextstoragesync.prod.cloudops.mozgcp.net"
},
{
  "rrname": 'prod.webextstoragesync.prod.cloudops.mozgcp.net',
  "rrtype": "A",
  "ttl": 1603,
  "rdata": "34.95.71.207" }
dns.rrname: webextensions.settings.services.mozilla.com
dns.foags: 8180
dns.rrtype: A
dns.type: answer
dns.qr: true
dns.rcode: NOERROR
dns.ra: true
```

Figure IV.24: Overview of data collected from a NIDS at the application level (source: personal work)

allow obtaining a parallel coordinate graph where the source and destination ports appear correctly, as shown in the figure IV.27. This real-time graph then clearly makes it possible to distinguish the source ports used by a machine (here, essentially due to numerous tabs in

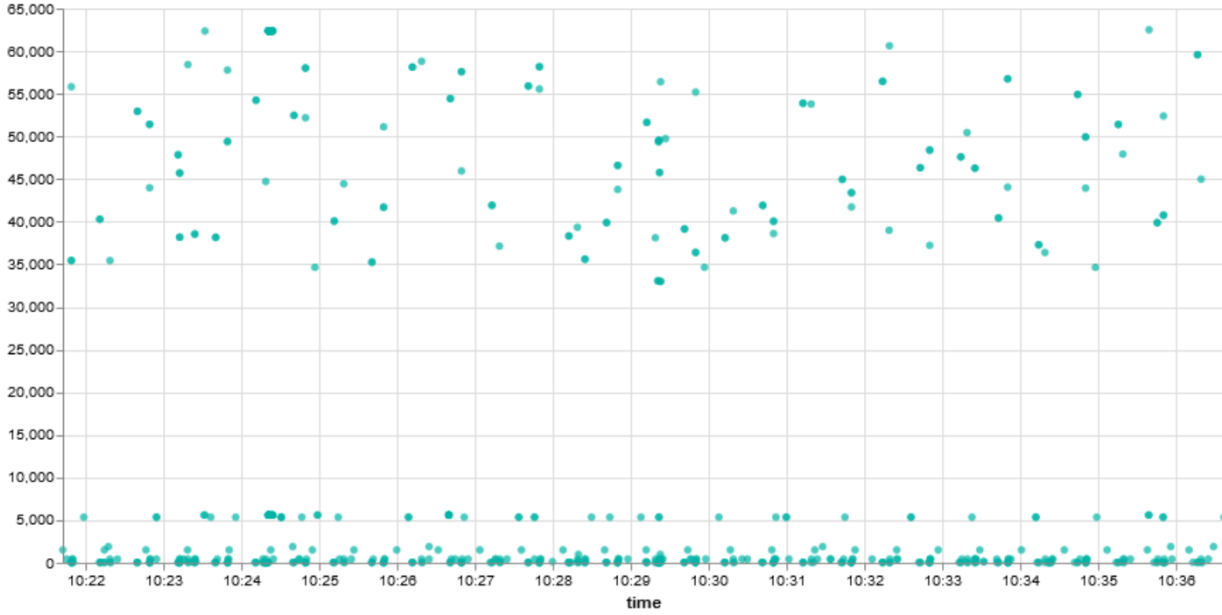


Figure IV.25: Graph of the *dest_port* metadata (ordinate) in time (source: personal work).

a web browser) and the destination ports (53, 80, 443). Produced over a longer period, this graph is representative of the data flows used within a network. In some cases (ship in a port of call with fiber connection, for example), the graphs obtained by using a compatible *netflow* tool remain however more complete, because they are independent of the capacities of the NIDS in terms of dissector. At the end, this type of graph allows, initially, to improve the KU of MIS and, moreover, facilitates the detection of anomalies in a graphical way.

IV.3.1 Dynamic risk assessment

Dynamic Risk Assessment (DRA) is an important challenge in cybersecurity research to carry out precise risk analysis and contribute to decision-making. Research on this subject is continuing [GGDM⁺18, Lag10, LPV13], including in a maritime context [PPM18]. Its realisation in real time would bring an important added value for the establishment of the different situation assessment processes necessary for the elaboration of the MCSA, but also to establish the links between the source information and of different formats. At the situation comprehension level, the DRA could facilitate data visualisation, analysis, and links between metadata. In terms of situation projection, the dynamic mapping of the constituents of a MIS would allow the identification of systems that could constitute potential targets due to their vulnerabilities. Finally, the DRA could also contribute to situation resolution

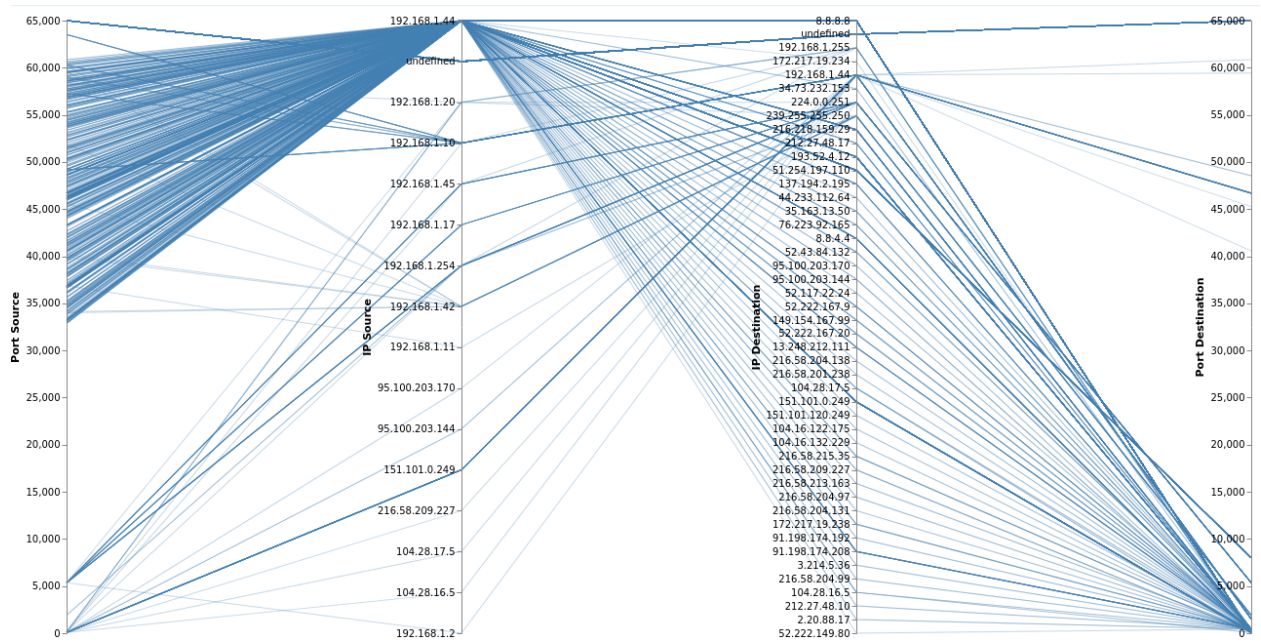


Figure IV.26: Realtime parallel coordinates graph without the proposed fix (source: personal work).

to contribute to decision making.

We wanted to contribute to two aspects to the DRA. On the one hand, we are interested in the visualisation of information of cyber interest based on graphs for the purposes of anomaly detection, in particular for data visualisation. On the other hand, we wanted to evaluate the added value of the DRA for decision making and cyber situation sharing. In the context of visualizing and understanding a large amount of data, decision making can be long and difficult. Indeed, in the event of detection, it requires being able to quickly identify the potential impacts (impact assessment), but also the cause (causality assessment). However, the data generally present in a SIEM are not always very well linked (we also speak of a pivot), and the operator must regularly switch between tools or databases to obtain a correct understanding of the situation.

Therefore, we have created a graph database providing the link between several types of data on our prototype Naval Cyber Range cyber monitoring platform. We use the enriched metadata, collected and indexed in the CE and make a link between them to ease the analysis (figure IV.28). We are also including the fix proposed in the IV.3.0.2 paragraph to ease connection tracking. The analyst of a SOC can then quickly identify suspicious connections from a machine or a network, using all the associated data present in the CE (time, wording of the alarm or logs concerned, resolved IP addresses, etc.). The contribution

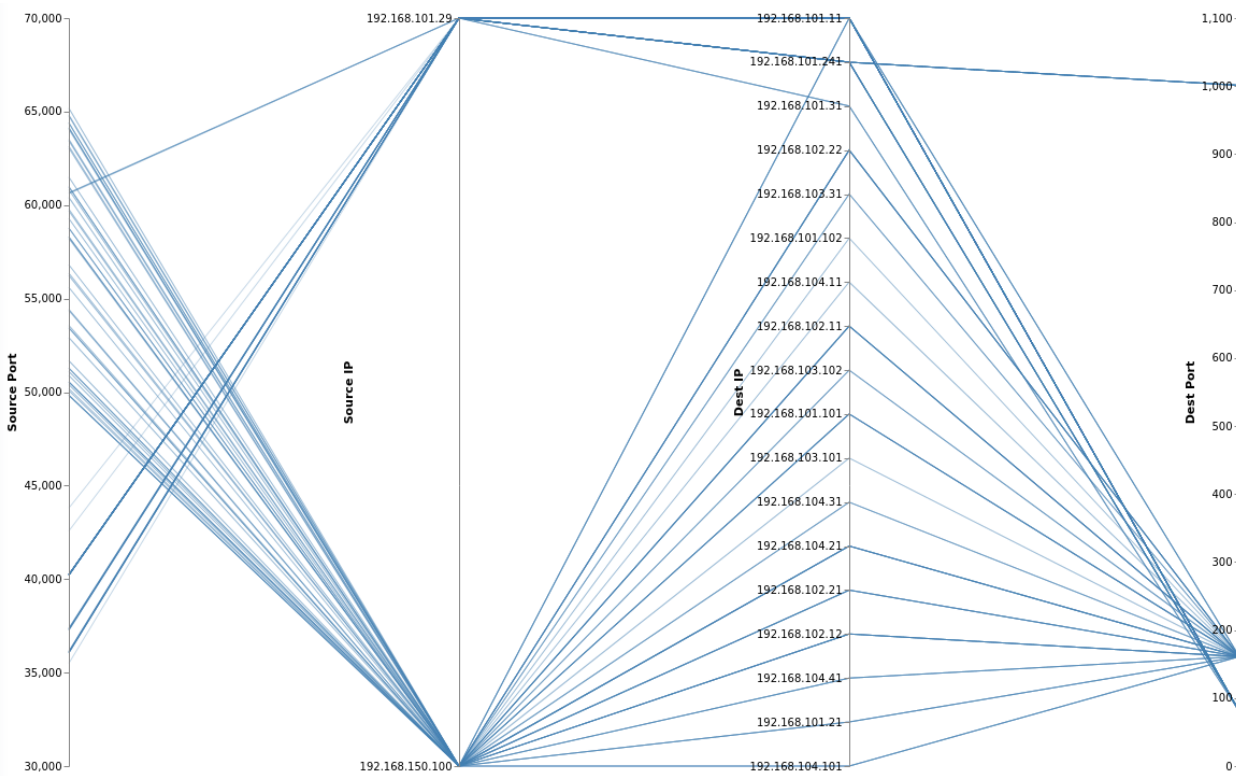


Figure IV.27: Realtime parallel coordinates graph with the proposed fix (source: personal work).

to the situation comprehension, causality assessment and situation resolution processes in the maritime context is therefore substantial.

In order to obtain the figures IV.29 and IV.30, we now link the data from the CE (metadata from on-board sensors) with the data from a CSIRT in order to instantly identify and graph the machines vulnerable to a particular Common Vulnerabilities and Exposures (CVE), in a passive manner (by the use of header metadata), or by using data sent by a HIDS. In the figure IV.29, we thus identify a machine (ip-12-34-56-7 ...) attacking the server "ship1_intranet" on January 7th, 2020, with a specific signature triggered by the sensor. The same graph (figure IV.30) also makes it possible to identify that the server in question has the Content Management System (CMS) "Wordpress" software installed, the version of which is vulnerable to a CVE, identified as CVE-2017-14723. The metadata of this CVE (Common Vulnerability Scoring System (CVSS), category, description) may suggest that it was exploited to lead an attack on the server. In a maritime context, this type of graph, which ensures a rapid fusion and link between data, greatly facilitates the identification of vulnerabilities, whether following the publication of new CVE, but also for analysis, by

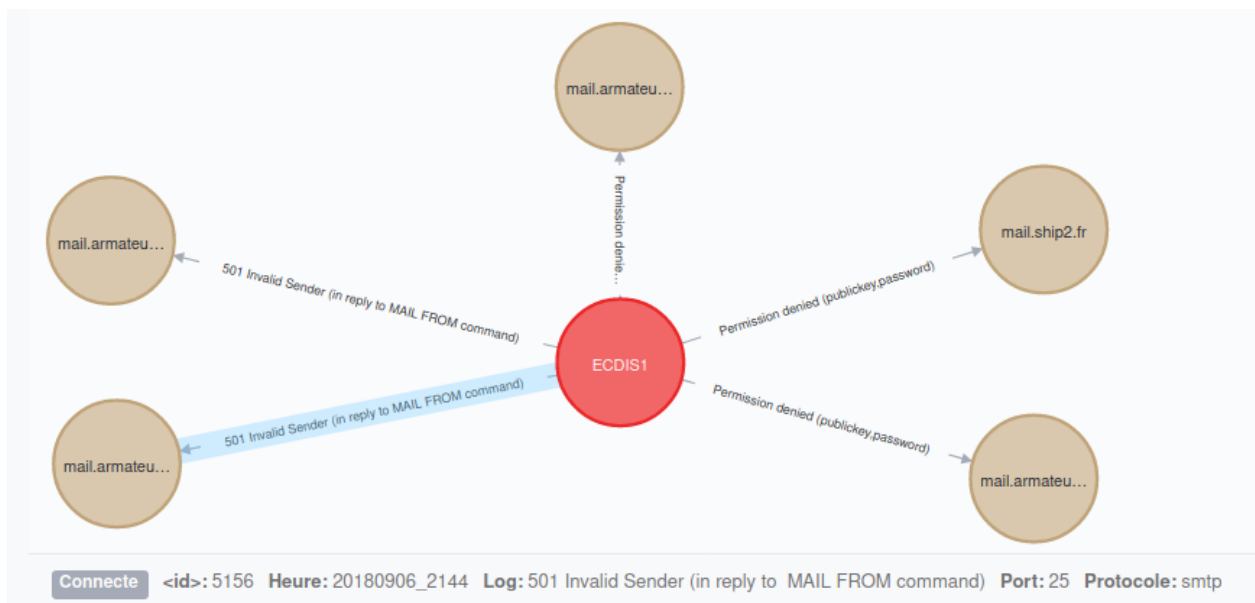


Figure IV.28: Use case of abnormal connexions from an ECDIS (source: personal work).

linking this information to potential security breaches.



Figure IV.29: visualisation and near real-time link of metadata related to an attack by the use of graphs (source: personal work).

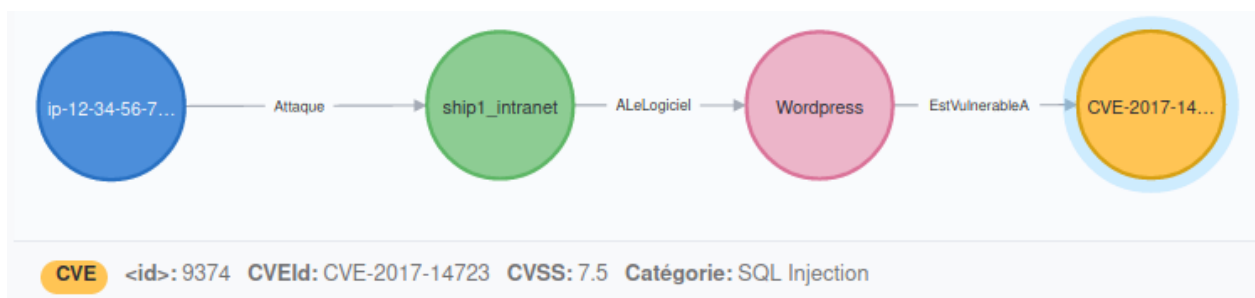


Figure IV.30: Added-value of the graph database to establish the real-time link between the signals of an attack, KU data and a potentially exploited CVE (source: personal work).

It should also be noted that this type of graph can contribute to the situation resolution.

In the figure IV.31, we import additional data related to the security patches issued by the vendors then create the link between the CVE and a possible patch from the vendor, specifying the corrected vulnerabilities, the patch download link, etc. Thus, between the detection of an alert (KT), the knowledge of the affected system (KU), the possible impacts, the CVE and the patch, all the processes necessary for the development of the MCSA is established and the saving of time and precision is particularly appreciable.



Figure IV.31: Added-value of a graph database from *situation perception* to *situation resolution* (source: personal work).

IV.4 Application of the MCSA by domain: detection and visualisation on NMEA 0183 data, contribution to the implementation of anomaly detection by machine learning

In this chapter, we end our work on RQ3 by addressing the issues of anomaly detection on MIS, and specifically on networks using NMEA 0183. NMEA 0183 is a proprietary standard developed from the 1990s, which allows the interoperability and interconnection of the various equipment used for ship bridges (ECDIS, AIS receiver or transponder, radars, sounders, various sensors, etc.). This standard allows the fusion and the representation of the information coming from the different sensors, generally at the ECS and ECDIS level. The analysis of the data of the network on which NMEA information passes can therefore make it possible to identify inconsistencies and anomalies in this information, the quality of which is essential for the development of the Maritime Situational Awareness (MSA).

IV.4.1 Context

Several studies and scientific articles have demonstrated the vulnerabilities of ECDIS as fusion tools [Svi19], but also of sensors that power them, such as AIS [BPW14, BG17,

Bur16] or GNSS such as the GPS⁶, sensitive to spoofing and jamming [ABPG20]. These vulnerabilities, which are becoming critical with the development of autonomous surface drones (Unmanned Surface Vehicle (USV))[Kav18], are regularly exploited by attackers in conflict zones and are grouped together under the term Navigation War (NAVWAR).

The fusion of information necessary for the development of the MSA on board a ship is therefore based on a standard, NMEA 0183, vulnerable by design and which collects multisensors data whose authenticity, sincerity and consistency cannot be guaranteed.

IV.4.2 Description of the NMEA 0183 standard

At the physical and electrical level, this protocol is based on the EIA RS-232 or EIA RS-422 standards to exchange data using the American Standard Code for Information Interchange (ASCII) encoding.

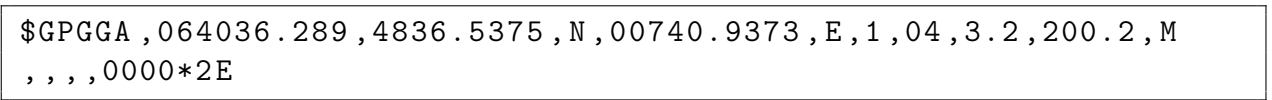


Figure IV.32: GPGGA frame sample on a NMEA 0183 network (source: [Wikipédia](#), license: CC BY-SA 3.0).

Figure IV.32 shows an example of a position frame transmitted by a GPS receiver on a network using the NMEA 0183 standard. The first two characters located after the \$ sign identify the equipment sending information (talker id). Thus, *GP* means that we are dealing with a GPS device, but other identifiers exist like *AI*, for an AIS device, *RA* for an Automatic Radar Plotting Aid (ARPA) type radar. The following three letters specify the type of frame: in this case, a *GGA* frame indicates the time of sending of the frame in universal time, the position in latitude and longitude, the type of positioning, the number of satellites used to compute coordinates, horizontal precision, altitude in meters, as well as several additional fields. The two characters after the "*" character correspond to a parity check (by an $\oplus$ operation on the set of characters included between the \$ and * signs). There are also other frames like *GPRMC*, giving latitude, longitude, date, speed and course over ground information, but without altitude information. We have identified more than a hundred references of talkers ids and 210 references of sentences ids, which we have

⁶<https://static1.squarespace.com/static/566ef8b4d8af107232d5358a/t/5c99488beb39314c45e782da/1553549492554/Above+Us+Only+Stars.pdf>

integrated, for those not already existent, in a NMEA dissector usable by network analysis software *tshark* and *Wireshark*⁷.

As part of our experimental platform, we set up a NMEA bus (versions 0183 and 2000) using a GPS Furuno ® GP-33 receiver and a AIS Furuno ® FA-30, as well as the associated antennas (*cf* figure IV.3). The NMEA bus connects each device to allow power and data exchange (transmission and/or reception). In our case, the GPS receiver transmits its data on an NMEA 0183 output to a multiplexer, made up of a Raspberry Pi equipped with a RS-422 *hat*. The data from the Raspberry Pi can then be sent with the User Datagram Protocol (UDP) protocol over a conventional Ethernet network and used to be displayed on the ECS. In reality, on a ship, many items of equipment can therefore be connected together, either on the NMEA bus itself, or after a multiplexer. If physical intrusion on a NMEA bus (addition of equipment, data injection) appears possible, we have carried out our work on the Ethernet connection and the NMEA data circulating there, because this point constitutes a network that is easier to access remotely and, therefore, more vulnerable. On a high tonnage vessel, this network is generally called Integrated Navigation System (INS), and is sometimes itself integrated into a larger network, which also includes on-board control/command systems and industrial network supervision and called IBS.

IV.4.3 Vulnerabilities of the NMEA 0183 standard

The NMEA 0183 standard is subject to numerous vulnerabilities which can be classified into two distinct categories: extrinsic vulnerabilities and intrinsic vulnerabilities. The first are linked to the sensors connected to the NMEA bus: they may themselves prove to be vulnerable, faulty, of poor quality, or subject to external disturbances (spoofing or jamming in the case of a GNSS) or even environmental: the quality of the data injected on the NMEA bus can therefore suffer, and little quality information is transmitted within the NMEA sentences themselves. The intrinsic vulnerabilities of the NMEA network are, for their part, linked to its design flaws (physical attacks, lack of a device that guarantees authenticity, integrity, confidentiality, availability, traceability and non-repudiation of data).

⁷<https://github.com/kmpm/wireshark-nmea/commit/5064cfcbb026d08454c392818ac6b64e42689881#diff-ed9edf58cclb06c05cca> and <https://github.com/kmpm/wireshark-nmea/commit/6fffd70c640961f2364b7ee966534437e144fc573#diff-ed9edf58cc775cca6a485eb01b4638e7>

IV.4.4 Peculiarities of the NMEA 0183 standard for anomaly detection

During our work on our prototype platform, we were able to make or obtain numerous captures of NMEA 0183 traffic before or after their passage through a multiplexer, the traffic then being transmitted using the UDP protocol and, to a lesser extent measure, TCP. Anomaly detection on a NMEA network can be done at three levels. The first, essentially after the traffic is multiplexed and encapsulated in the UDP protocol, is to detect anomalies in the IP traffic flowing on the INS network. The second comes down to analyzing the envelope of the NMEA frame and its characteristics. Finally, the third level concerns the dynamic data transmitted by the standard (for example: geographical positions, AIS tracks, environmental data, etc.), sometimes called "payload".

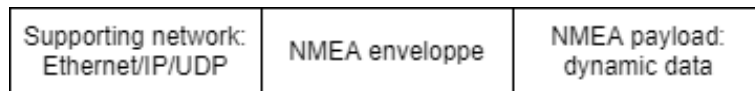


Figure IV.33: The three possible levels of anomaly detection on a NMEA 0183 network, once multiplexed and encapsulated (source: personal work).

IV.4.4.1 Peculiarities of the NMEA 0183 envelope

From these captures, we isolated three significant captures of NMEA traffic, named *cap1*, *cap2* and *cap3*. The IV.3 table summarizes the characteristics of these captures, which represent a total of 9.3 million packets. Two remarks can be made in the first analysis: the average size of the packets remains stable and low. On the other hand, the average speed and the number of packets vary: in a deterministic *a priori* network, we hypothesize that this variation is linked to AIS and radar data, a hypothesis confirmed by isolating the concerned network traffic. This difference is explained by the fact that, depending on the position of the ship (at quay, at sea), its navigation zone, the time of day, or even the ionospheric propagation conditions, the maritime traffic captured in the environment of the ship varies greatly: the volume of data re-transmitted on the NMEA network by the AIS and radar sensors is therefore likely to vary significantly.

Table IV.3: Characteristics of the NMEA traffic capture files

Name	Number of packets	Capture duration (s)	Packets/s	Average packet size	Average throughput (b/s)
cap1.pcapng	22 707	385	58,9	86	5044
cap2.pcapng	22 574	209	122,3	87	10 651
cap3.pcapng	29 330	234	125,3	87	10 889

IV.4.4.2 Analysis of the NMEA 0183 standard envelope

In this subsection, we carry out an analysis of the envelope of the NMEA 0183 standard. The objective of this analysis is to allow features identification. By design, we hypothesize that the NMEA traffic must have rather deterministic characteristics, in particular because the sensors probably transmit their data at regular intervals. We are therefore looking for the characteristics of a deterministic system on the envelope of NMEA frames. Figure IV.34 thus indicates the distribution, in number, of the phrases PHKS (proprietary), AIVDM (AIS), RATTM (RADAR) and PHTRO (proprietary), which all have a deterministic and significant volume, while the NW talker sentences have an equal distribution in quantity (4.21% of captured frames). Finally, the frames related to positioning and environmental parameters (GCGGA, GCVTG, GCZDA, GPRMC, NAVHW, SDDBT, SDDPT, SPDPT) all represent a distribution equal to 2.16%.

These characteristics are confirmed when we analyze the distribution of the packets of each sentence over time: if some distributions do not display a deterministic characteristic, because they depend on extrinsic factors⁸, for others, the distribution over time appears particularly precise and regular (appendix E). This deterministic characteristics, which seems to be distinguished, makes it possible to propose examples of particular measurements for these different sentences and the possible selection of features.

We analyze these files in Table IV.4 for sentences that display a deterministic profile (*ie* all except AIVDM and proprietary PH and IO sentences). We note that several frequencies can be isolated: 1 Hz for the sentences coming from the RA talker (OSD and RSD), 2 Hz for

⁸For instance, the number of ships detected by radar or by AIS for the RATTM and AIVDM sentences.

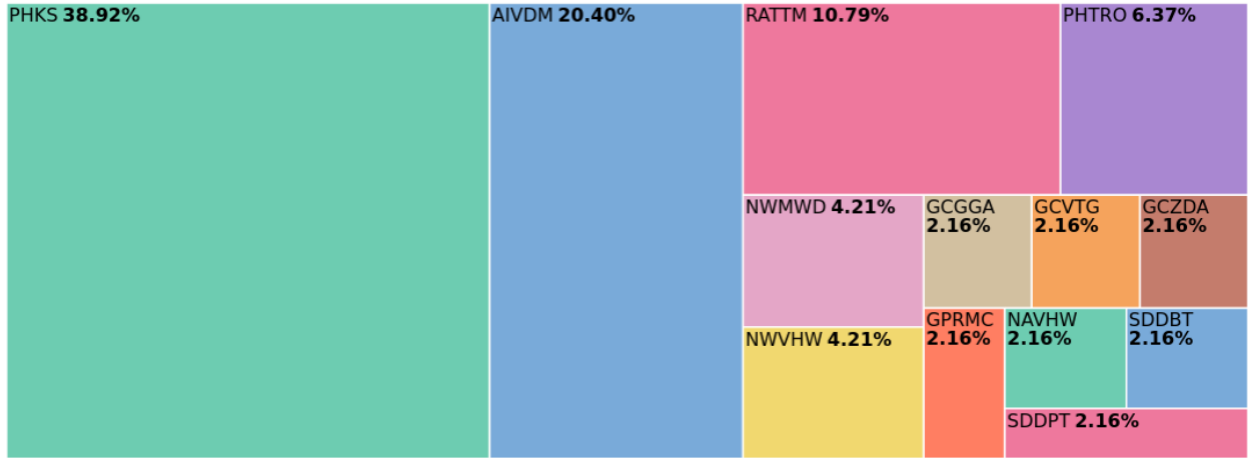


Figure IV.34: Repartition, in quantity, of the type of sentences IDs on the files *cap1.pcapng*, *cap2.pcapng* and *cap3.pcapng* (source: personal work).

the talkers GC, GP, NA, SD and SP, 4 Hz for the NW talker and 10 Hz for the RATTM sentence RATTM, the approximations being due to the capture window, the capture technology and the accuracy of the clock on the computer conducting the capture.

This determinism in time could thus make it possible to detect, over a given duration, a regular cycle of frames specific to the given system and facilitate the detection of anomalies, for example by using the concept of Cybernetic-DNA, which could apply with interest to a deterministic network like this. [MK19]

On the captured files, we have identified that some sentences always have the same size (Table IV.5). Other frames have variable sizes, such as the RATTM, GGA, GLL, TXT or even VTG sentences.

The captures which were conducted also make it possible to identify several cardinalities. Thus, the number of talkers and the associated sentences are known and do not change over time.

We also notice that some NMEA sentences are emitted by several talkers. It should be noted that, if the characteristics of the NMEA network are linked to sensors (and can therefore be modified), they can also be transformed by intermediate devices, such as multiplexers: the characteristics identified are then specific to the multiplexer, and not to the analysed sensor. Following our work on the talker and sentence ids, we also hypothesised that it was possible, from the capture files, to reconstitute the INS network at the origin of the traffic.

Table IV.4: Characteristics in period and frequency

Talker ID, Sentence ID	Average period (s)	Average frequency (Hz)
GCGGA	0,497	2,00
GCRMC	0,497	2,00
GCVTG	0,497	2,00
GCZDA	0,5	2,00
GPGGA	0,497	2,00
GPRMC	0,497	2,00
NAGGA	0,497	2,00
NAVHW	0,497	2,00
NAVTD	0,497	2,00
NWGLL	0,256	3,90
NWMWD	0,256	3,90
NWMWV	0,256	3,90
NWVHW	0,256	3,90
RAOSD	0,991	1,00
RARSD	1,00	1,00
RATTM	0,099	10,02
SDDBT	0,497	2,00
SDDPT	0,497	2,00
SPDBT	0,497	2,00
SPDPT	0,497	2,00

Starting from new capture files, we are able to model the functional architecture of the underlying NMEA network (Figure IV.35). We represent the themes of the NMEA sentences in different colors: black for AIS, blue for positioning and time information, orange for the proprietary information, yellow for navigation information, red for radar, green for environmental information.

This modeling makes it possible to emit several hypotheses. First of all, certain devices originating from the NMEA sentences (the four IP addresses to the left of the diagram) are directly connected to the network. This is the case with the 192.168.0.2 device, which can be identified as an ARPA type radar. This radar emits, in addition to RATTM and RARSD type sentences, proprietary sentences (IOTXT) which could allow it to be identified with more precision (brand, or even type) but, according to the references found in open

Table IV.5: Examples of characteristic sentence sizes

<i>Sentence IDs</i>	Frame size (bytes)
DPT	70
MWD	77
MWV	71
NHW	76
OSD	79
RMC	121
RSD	93
TRO	71
VDM	91
VHW	76
VTG	86
ZDA	81

Table IV.6: Talkers and sentence IDs of the *cap1*, *cap2* and *cap3* files

Equipment	<i>Talker ID</i>	<i>Sentence ID</i>
192.168.0.1	NW, PH, SD, SP	TRO, MWD, VTG, DPT, DBT, VHW
192.168.0.2	RA, IO	TTM, OSD, RSD, TXT
192.168.0.3	PH, AI, GP, NA, NA, PH	KS, VDM, GGA, VHW, VTG, TRO, RMC

source, there are several manufacturers whose radars emit this type of sentences. The direct connection to the network can also be confirmed for the 192.168.0.4 equipment, which has all the characteristics of a GNSS receiver: it is probably the second GPS receiver of the vessel in question, for which the talker id has been changed from GP to GC to allow separation of the origin of positions between the two receivers.

Then, devices with IP addresses 192.168.0.1 and 192.168.0.3 are, for their part, the source of multiple sentences of different talkers. They are, therefore, most likely, multiplexers or software or hardware performing these functions. For the 192.168.0.1 device, it combines ECDIS functions and receives then re-transmits information related to environ-

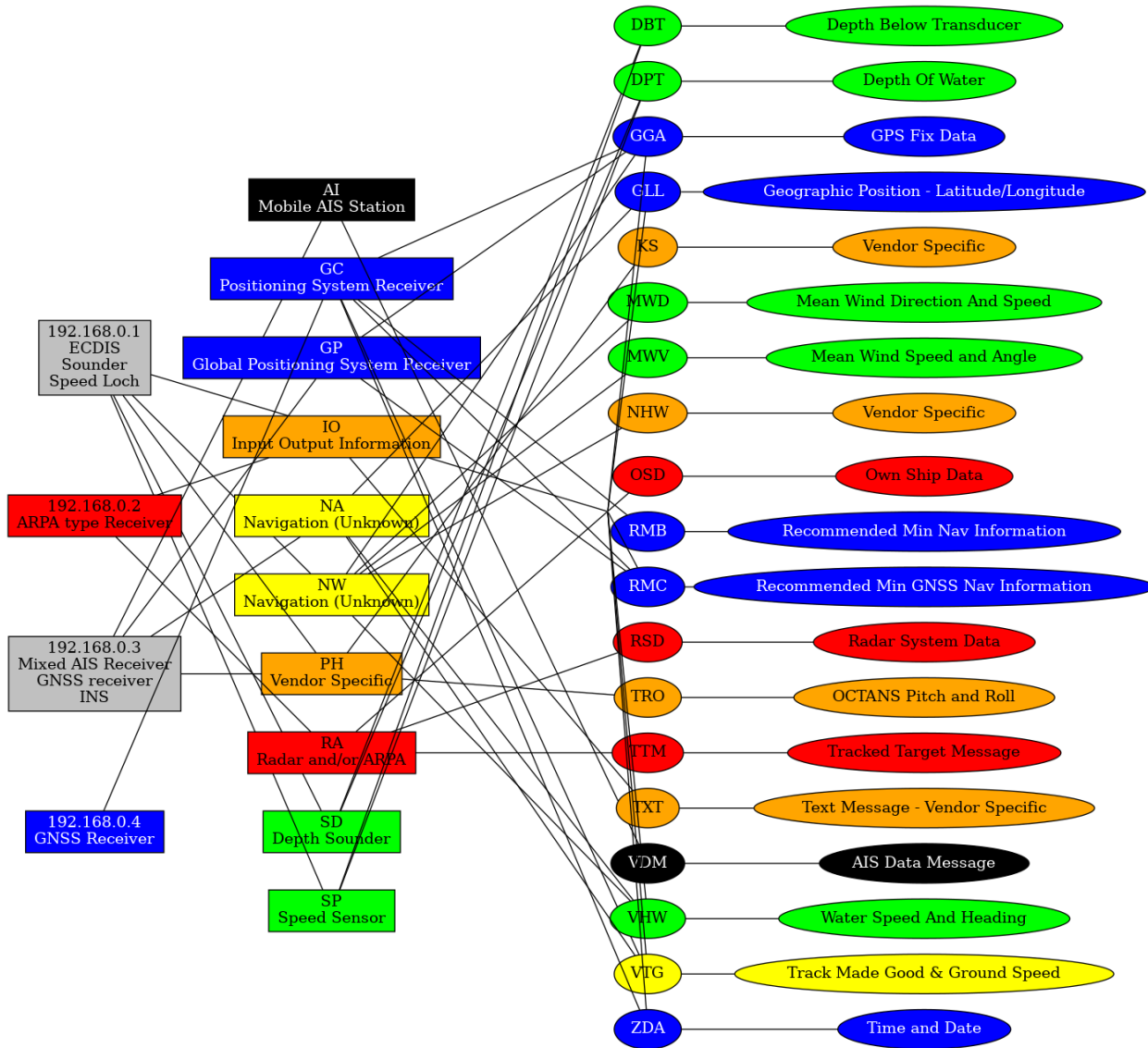


Figure IV.35: Model of the NMEA network, talkers and sentences IDs from capture files (source: personal work).

mental parameters (speed, water height). For the 192.168.0.3 equipment, it is probably an AIS receiver coupled with a GNSS receiver and an inertial unit or roll/pitch sensors.

The network can then be modeled as in figure IV.36.

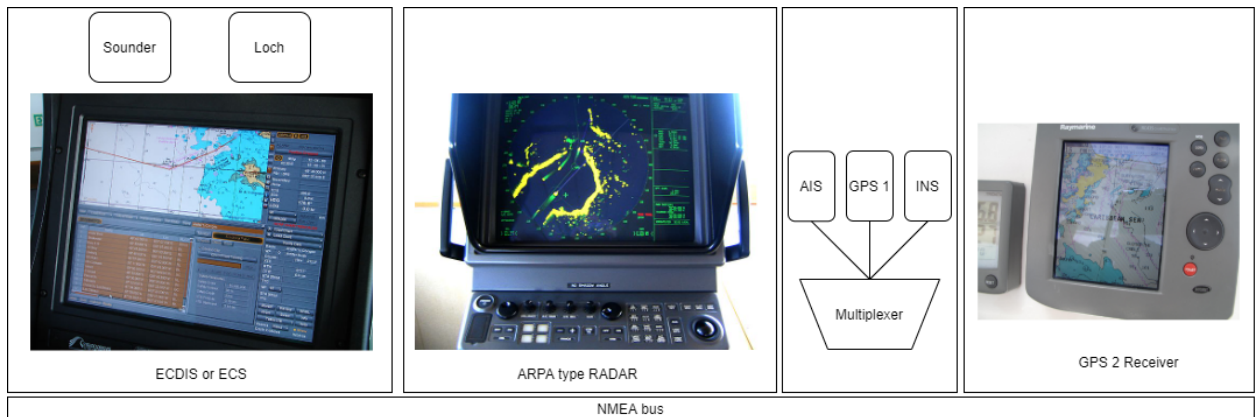


Figure IV.36: Model of the overall NMEA architecture from files *cap1*, *cap2* and *cap3* (sources: ECDIS: [Wikimedia Commons](#), unmodified picture, author: [Hervé Cozanet](#), license: [CC BY-SA 3.0](#), RADAR: [Wikimedia Commons](#), unmodified picture, author: [Clipper](#), license: [CC BY-SA 2.5](#), GPS: [Wikimedia Commons](#), unmodified picture, author: [Fairley](#), license: [CC BY-SA 2.0](#)).

IV.4.5 Results

To detect the exploitation of an intrinsic vulnerability (injection of spoofed frames), the anomaly detection on the transport network is generally carried out by conventional intrusion detection systems. Given the significant consequences and the occurrence of recent GNSS spoofing and jamming incidents, we have chosen to focus our work on the detection of anomalies on the NMEA 0183 payload of GNSS data (extrinsic vulnerabilities). We have implemented, within the Naval Cyber Range, a GPS spoofing tool and configured it so that it can be effective on maritime GPS equipment. These spoofing operations were carried out in the laboratory, at very low power, using a direct connection between the transmitter and the receiver by means of a coaxial cable and without emission in the radio frequency space. Our final goal is that, for each spoofing event, the sensor detects anomalies and informs the operator via the ECS.

In fact, when anomalies are detected, it appears essential that the operator of the installation be informed in a timely manner, in order to carry out the necessary actions to stop the attack and, if possible, that he be helped and supported in his *situation resolution* actions. The determination of the *CoA* can be based on the knowledge of the attacker (TTP) to anticipate his next actions and contribute to the situation projection, but it is also very dependent the type of information system and existing resilience (either technical or organisational). This subject is the scope of several research works, in particular within the

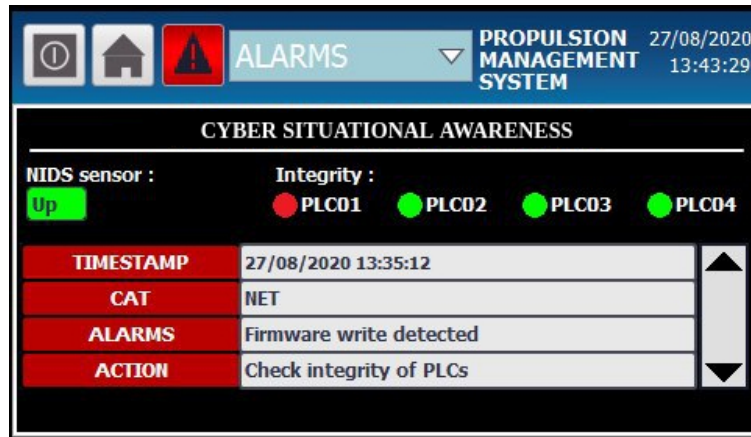


Figure IV.37: Information sharing and enhancement of the Maritime Cyber Situational Awareness using a control and alert panel on an industrial system of the Naval Cyber Range (source: personal work).

framework of the CACAO project⁹.

As part of our work on our prototype platform, we looked for the implementation of a solution at the tactical level, as it is generally absent from field equipment. This is particularly the case on OT systems, when the cyber detection information goes back to the SOC, but is not shared at the level of the operator, which however seems often the most appropriate to receive the alert in order to see if the possible cyber alarm could have had physical consequences. The development of the prototype within the framework of our work allows a concrete contribution at the tactical level, by allowing the installation of control panels and dedicated warning directly on the industrial systems and the MIS. Thus, Figure IV.37 presents the information panel set up on the racks of the industrial systems racks on the Naval Cyber Range and which allows the operator to quickly visualize the state of the sensor, the integrity of the programs of the various PLCs, but also the alarms detected and the proposed solutions.

We also worked on the NMEA standard itself, in order to be able to share the detection information of the cyber sensor. We therefore propose an evolution of the NMEA 0183 standard to include information related to cybersecurity within the framework of a project called Cyber Situational Awareness Made Easy for NMEA (CARMEN). The concept is based on the creation of a new NMEA talker called CY, corresponding to the detection sensor. For each feared event of the NMEA standard, we then created a specific sentence identifier (*cf* table IV.7). In the long term, the CARMEN project will include all of the

⁹https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=cacao#overview

Table IV.7: Examples of cyber talker ID and sentence ID cyber and suggested COA within the CARMEN project

<i>Talker</i>	<i>Sentence</i>	Description	Suggested COA
CY	AIJ	Cyber / AIS Jamming	<i>Strengthen visual surface awareness</i>
CY	AIS	Cyber / AIS Spoofing	<i>Strengthen visual surface awareness</i>
CY	GPJ	Cyber / GPS Jamming	<i>Use alternative positioning solutions</i>
CY	GPN	Cyber / GPS Nav Alert	<i>Check GPS nav information</i>
CY	GPS	Cyber / GPS Spoofing	<i>Use backup positioning solutions</i>
CY	GPT	Cyber / GPS Timing Alert	<i>Check GPS timing information</i>
CY	NET	Cyber / Network intrusion	<i>Check NMEA network integrity</i>
CY	WDG	Cyber / Watchdog	<i>Check cyber sensor function</i>

feared events depending on the sensors and actuators present on a NMEA network and for which a cyber sensor could verify the consistency.

For instance, in the case of Figure IV.38, we exploit an intrinsic vulnerability of the NMEA network which makes it possible to intercept NMEA frames and inject new ones in order to distort the display of the position on the ECS. On our prototype platform, in the absence of an anomaly, and if it is functioning correctly, the cyber sensor constantly sends a good functioning indication frame (\$CYWDG). This information is relayed to the ECS which indicates that the sensor is working correctly and that there are no anomalies. If an anomaly is detected on the sensor (in the event of spoofing: non respect of determinism, unknown transmitter address, inconsistent position with respect to the kinematics, etc.), the probe emits a \$CYGPS type alert and the plugin developed on the ECS displays an alert (red dot) and suggests a workaround (Course of Action, for example: use an alternative positioning method).

In the case of Figure IV.39, we exploit an extrinsic vulnerability of the NMEA network (spoofing of the GPS sensor), which allows the injection of bad position information (or time) in order to distort the display of the position or the operation on the ECS. During the spoofing operation, carried out by the Software Defined Radio (SDR) transmitter, the sensor detects the anomaly (in the event of a spoofing attempt: change of information linked to the constellation, position inconsistent with respect to the kinematics, strong variation in weather, etc.) and emits a \$CYGPS type alert and the plugin developed on the ECS displays a simplified alert (red dot) and suggests a workaround (Course of Action, for example: use

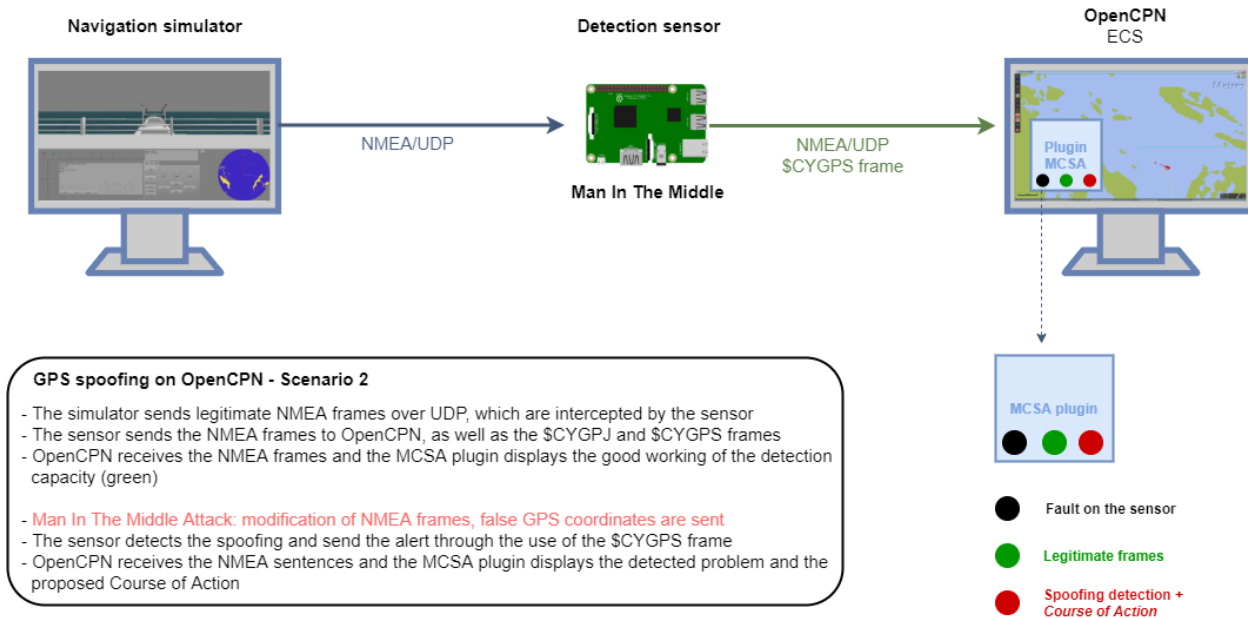


Figure IV.38: Exploitation of an intrinsic vulnerability of the NMEA protocol and alerting scheme to the ECS to display the appropriate CoA (source: [Naval Cyber Defence Chair](#), reproduced with the kind permission of the author).

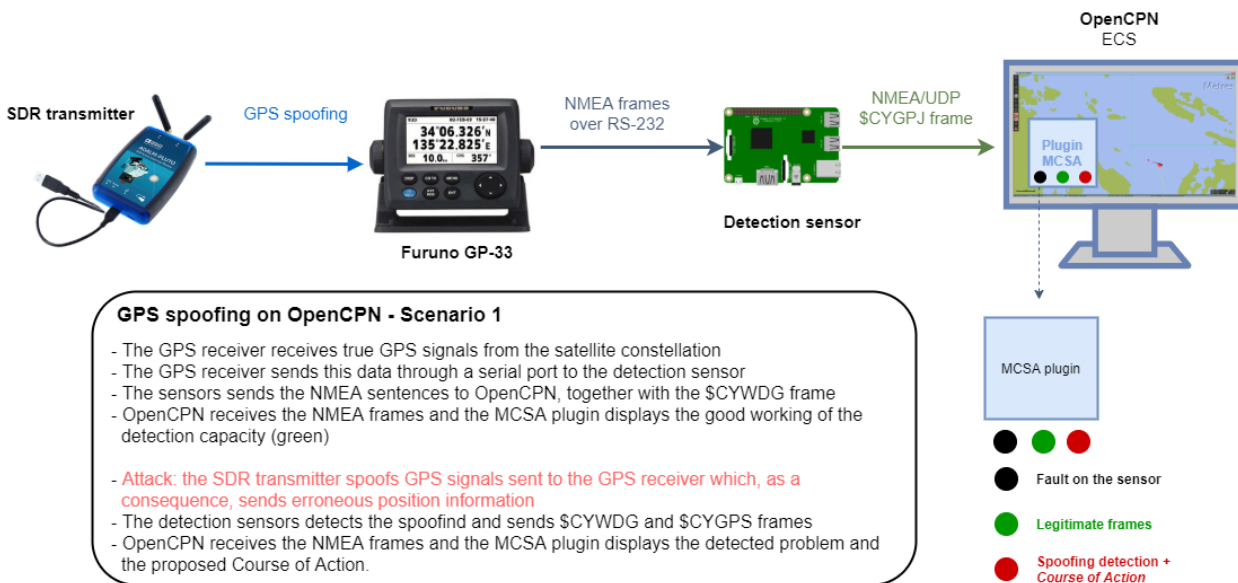


Figure IV.39: Exploitation of an extrinsic vulnerability of the GPS network and alerting scheme to the ECS to display the appropriate CoA (source: [Naval Cyber Defence Chair](#), reproduced with the kind permission of the author).

an alternative means of positioning).

At the ECS level, in the absence of the GPS spoofing, the CARMEN prototype indicates

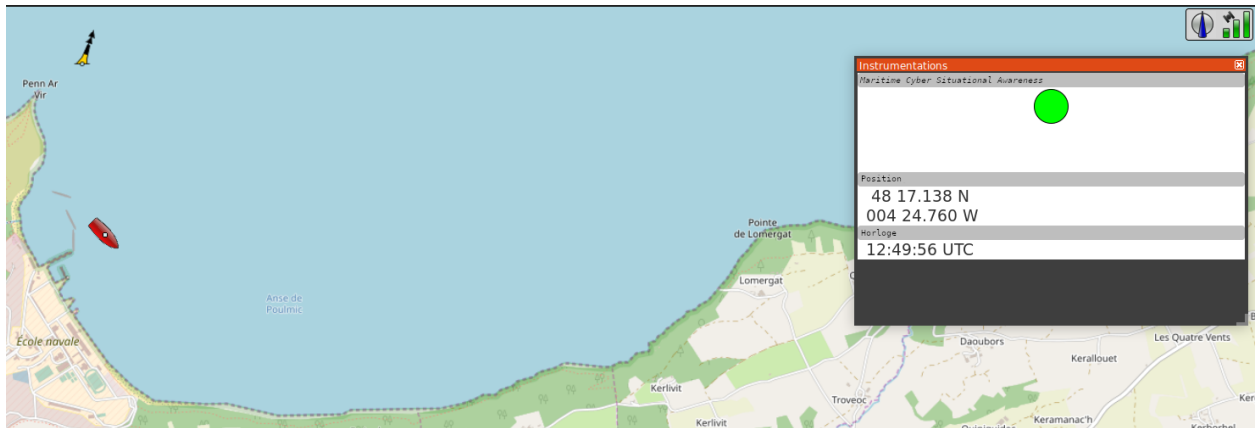


Figure IV.40: Display of the CARMEN plugin in the ECS in the absence of spoofing (source: [Naval Cyber Defence Chair](#), reproduced with the kind permission of the author).

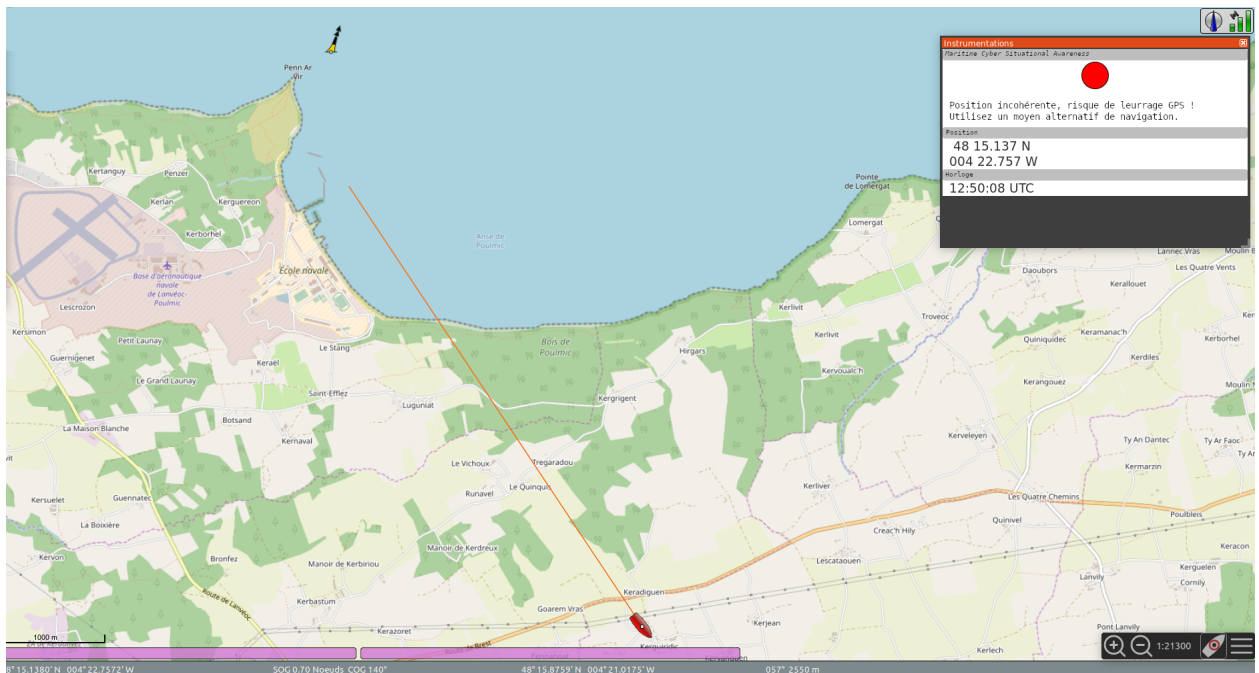


Figure IV.41: Display of the CARMEN plugin in the ECS during spoofing, with the suggested CoA (source: [Naval Cyber Defence Chair](#), reproduced with the kind permission of the author).

that the link with the sensor is functional and that no spoofing is detected (Figure IV.40).

We then perform a GPS spoofing with a significant position shift: the CARMEN prototype immediately indicates the detection and offers an alternative means (Figure IV.41). The spoofing is deliberately important so that the position shift is particularly visible: the quality of detection is, in reality, much finer.

IV.5 Conclusion

In this chapter, we have dealt with the research questions RQ1 (collection and modeling architecture), RQ2 (data analysis and representation) and RQ3 (anomaly detection on a maritime information system). Our experiments on a prototype platform have led us to integrate into the Naval Academy's Naval Cyber Range several essential components, in terms of anomaly detection, visualisation and information sharing on maritime protocols. This integration was notably carried out through an application architecture which allows the elaboration of the Maritime Cyber Situational Awareness.

In order to validate our work, we have proposed a set of quality measurement criteria which make it possible to ensure that the MCSA elaboration is carried out with a high level of confidence in the collected data to avoid any bad interpretation throughout the perception, comprehension, projection and resolution phases.

Subsequently, we worked on the platform to create dynamic dashboards and on the source codes of a network analyzer and a NIDS, in order to improve the accuracy of the data that they generate and resulting graphs. The gain in precision makes it easier to understand the situation and make decisions, while also improving the quality of the mapping of maritime information systems.

The contributions of Maritime Cyber Situational Awareness to the Dynamic Risk Assessment concept were then demonstrated. The use of graph-oriented databases and the addition of additional data from CSIRT have made it possible to underline the added value of these tools for SOC analysts: these works, inconceivable without a architecture tailored to the maritime world, made it possible to create an effective link between all the phases of the MCSA, from detection to resolution for the operational and tactical levels.

Finally, the work carried out on an experimental platform made it possible to demonstrate the added value of an adapted and efficient cyber monitoring architecture for the purpose of detecting anomalies on maritime information systems. For this, research focused on the characteristics and vulnerabilities of the NMEA 0183 standard and satellite positioning systems, in particular by looking for deterministic characteristics and attributes relevant for the detection of anomalies by machine learning.

The lack of operator information in the event of a cyber anomaly, frequently encountered on OT systems, motivated the development of the CARMEN project. This project made it possible to propose an evolution of the NMEA 0183 standard to include exchanges of cyber

intelligence between a detection sensor and an ECS or an ECDIS. During a prototyping phase on the Naval Cyber Range platform, the detection of anomalies, the understanding of the event and its resolution were the subject of instrumentation in the form of a prototype and presented in a simple, yet comprehensive manner, to the operator on industrial and navigation systems.

By forwarding information from Maritime Cyber Situational Awareness to the global Maritime Situational Awareness tool made up of ECS or ECDIS, the interest of fusion between SAs is demonstrated. The overall state of knowledge offered to the officer of the watch and the ship's captain allows them to establish a link, if necessary, between cyber anomaly and surface situation or ship's integrity, while helping him to make the most appropriate decision.

General conclusion and perspectives

V.1 Problematic

Modern ships can be viewed as complex systems of systems, increasingly digitized, with IT, OT and cyber physical systems essential to the conduct of their mission. They exploit vulnerable hardware, software and protocols that can weaken them and put them within reach of compromise by a motivated attacker. The perception and understanding of a cyber event, its projection in time and its resolution, which are current fields of research, become all the more complex when applied to the maritime context and to a mobile object such as a ship.

We proposed to adapt and enrich an existing concept, called situational awareness, to cybersecurity in the maritime domain by declining it under the name of Maritime Cyber Situational Awareness. The objective of this approach was to apply a methodology successfully followed in other fields, in particular in aeronautics, to maritime cybersecurity, not by directly using tools such as Head-Up Display, but rather for the relevance of its founding principles. The direct application of this concept to the maritime world is proving difficult, because the constraints specific to the environment do not allow it to be monitored like a conventional IT information system.

We have thus identified three research questions: data collect architecture and modeling (RQ1), data analysis and representation (RQ2) and anomaly detection within the maritime context (RQ3).

V.2 Results

During our work, we wanted to demonstrate that, in order to respond appropriately to the problem of anomaly detection in the maritime community, it was first of all necessary to know and understand the sector as a whole. This is the reason why we started by carrying out a global maritime community analysis, and more particularly of ships, their constraints and their systemic vulnerabilities in terms of cybersecurity. It is on the strength of these observations that the concept of Maritime Cyber Situational Awareness was then developed, by adapting and declining the principle of situational awareness already used in the military community.

Subsequently, our work on RQ1 led us to define the potential global architecture that could prove to be effective in ensuring the cyber monitoring of a ship and a fleet of civilian ships, both for the onboard and the shore part. We have shown that this architecture is feasible and can be qualitative, provided that it is adapted to the peculiarities of the maritime environment. In order to validate the level of quality reached, precise criteria were defined and then implemented on a prototype. They made it possible to demonstrate the efficiency of the architecture, even in the case of specific constraints such as the degradation of the support network.

Data analysis and visualisation remain an important research issue for cybersecurity, in a context of heterogeneous data and large volumes. To respond to RQ2, modern and advanced tools have been chosen, such as graph databases and big data processing technologies, in order to provide concrete and rapid answers to questions that arise in terms of cybersecurity on complex and interconnected systems of systems. We have also contributed to the improvement of two important tools for data analysis and visualisation, namely an intrusion detection sensor and a network analysis tool.

To answer to RQ3, work on the NMEA standard and in particular the CARMEN project has made it possible to highlight the peculiarities and vulnerabilities of the standard. Faced with the increasing number of GNSS jamming and spoofing cases, research is faced with an insufficient number of datasets. A prototype was therefore developed in order to generate the necessary data sets for our research work for anomaly detection in position, navigation and time, which are essential information for a ship. To follow up on the problem raised by this research question, our work then focused on the sharing of information to monitored systems, which is regularly lacking. Thus, the CARMEN experiment on a network sharing GNSS and ECS data on the Naval Cyber Range made it possible to display the detection information

directly on the Human Machine Interface (HMI) of an ECS, in a format understandable by the user, thus providing a proof of concept of the fusion of Maritime Cyber Situational Awareness and Maritime Situational Awareness.

Finally, beyond the contributions for the maritime world, the work on the Naval Cyber Range has endowed the French Naval Academy and, more broadly, the French Navy, with a realistic tool and, at present, still unique in Europe, which opens up many research perspectives in the detection of anomalies, but also in the protection of maritime systems.

V.3 Discussion

The MCSA concept largely responds to the cybersecurity issues raised for the sector. It does, however, have limitations. In this section, we want to engage in open discussion and reflection on the limitations of our work.

First of all, the bias of carrying out a rather holistic research on cybersecurity for the maritime community must be explained. When carrying out the state of the art of our thesis, it was noted that the number of scientific references relating to maritime cybersecurity was limited, in particular with regard to the analysis of the systemic causes of vulnerabilities in maritime information systems: devoting part of our work to it therefore seemed necessary to us.

Next, the quality of cyber monitoring relies primarily on sensors. In the current state of research and especially the software solutions proposed, the choice remains relatively limited to sensors using signatures or behavioral approaches. The two solutions each have their performance and limitations in terms of detection quality. Rather than carrying out a new study on the performance or the shortcomings of one or the other, our approach to using network or host sensors had the main objectives of contributing, in addition to anomaly detection, to improving knowledge of systems (KU), which is often too incomplete in the maritime community.

The architecture that we have chosen and tested, in particular on the information sharing part, focused on the tactical level. This choice is explained by the fact that sharing at the tactical level is probably the most efficient in terms of reaction and correlation on a cyber physical system, but also because the prototype used on the Naval Cyber Range is of tactical level. However, a important work remains to be carried out at the "high" level

even if, very often, they can go beyond the strict maritime framework (such as information sharing with a national cybersecurity authority, for instance).

The quality criteria retained and measured during our work are particularly linked to the Naval Cyber Range and are not intended to be exhaustive or strictly representative of the reality. It lacks, in particular, the representation of vessels of different types with specific information systems, or even a complete fleet. However, this generic architecture has enabled us on several occasions to resolve anomalies or doubts on the operation of certain sensors, on their updating, as well as on the operation of the support networks. Therefore, this generic architecture model deserves to be continued. Other quality measurement criteria or sub-criteria could probably be added later.

Much research has already been carried out on the issues of visualisation and its use for anomaly detection. This is the reason why we wanted to focus our efforts in the directions which seemed the most relevant and realistic to us. This was the case for parallel coordinate graphs, but also and above all for dynamic risk assessment. This concept, which is sometimes difficult to model, takes on its full meaning in the maritime context, in particular thanks to the use of graph databases which greatly facilitate its implementation to make it easily achievable.

Anomaly detection work on the NMEA 0183 standard has brought out several deterministic attributes that can be used for anomaly detection purposes based on machine learning techniques. Due to the volume and the heterogeneity of the data and sensors, but also because the subject partly diverged from the initial theme of the thesis and would itself require a full-fledged thesis, not all avenues could be explored, in particular as regards the deterministic character of certain values.

Finally, the work carried out in sections IV.4.5 and IV.5 within the framework of the CARMEN project and in particular the proposals of Course of Action, would benefit from being enriched. It should however be remembered that the visualisation on the ECS, the ECDIS or on the industrial system is primarily intended for a non-expert operator, the expert rather having recourse to the information provided, for example, in the LE and analysed by the SIEM or a graph-oriented database.

V.4 Perspectives

In Figure V.1, we offer a certain number of perspectives for research and development of our work on the concept of Maritime Cyber Situational Awareness, on the data, the model in itself and on its overall architecture. While some perspectives have already been taken into account within our research chair, where several future theses will benefit from the areas of research and development identified, other research perspectives remain to be initiated.

V.4.1 Data

Among the points that we identify, let us first of all evoke the capture of mass data and the associated analytical capacities, which should make it possible to broaden the use of the solutions mentioned and to exploit the data for purposes other than strictly cyber: we can think of hardware anomaly detection on cyber physical systems (D1) or use-cases for operational detection (D2).

The work related to the quality of the data collected from sensors can be continued: the modeling and creation of the Naval Cyber Range now make it possible to have a large quantity of realistic data, which will benefit from being valued and made available for the world of research and education (D3).

The demonstrations carried out on the topic of Dynamic Risk Assessment herald research of great interest. The MCSA and the implemented architecture now allow to achieve relevant results. Experiments should be continued in order to validate the added value of DRA in the multiple and complex contexts of the maritime community (D4).

The chronological follow-up of the transmission of NMEA frames over a largely deterministic network such as a INS could be the subject of an interesting analysis using signal processing or Cybernetic-DNA techniques, in order to identify patterns of time distribution of sentences and talker IDs and to strengthen anomaly detection by an original approach. Anomaly detection on the payload and in particular on dynamic data (position and navigation), but also on time, will require additional work (D5).

The fusion and correlation work of the MCSA with other SAs (security, safety ...) will deserve to be continued. This fusion of data from different sensors will further improve the fineness of detection. Thus, while limiting the number of false positives, the coupling of data from radio-frequency (RF) and cyber sensors will make it possible to obtain data of a high

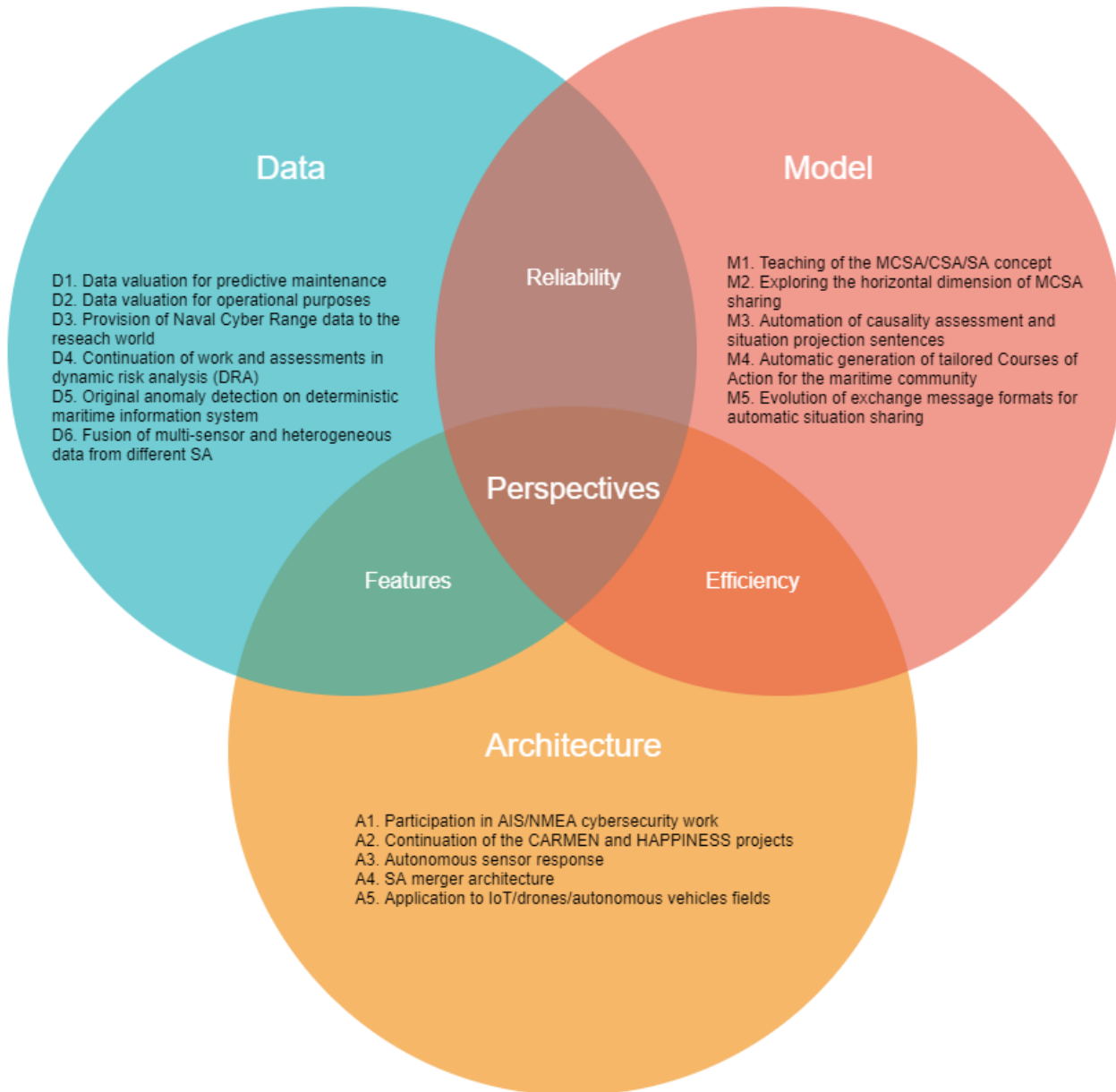


Figure V.1: Research and development perspectives for the Maritime Cyber Situational Awareness concept (source: personal work).

level of completeness and precision, in the face of a significant threat to the radio-electric spectrum. The same logic could be followed to correlate alerts from a cyber sensor (*eg*: detection of a network anomaly) with the incoherent evolution of a cyber physical sensor or a radio-electric sensor, for example by the use of a scoring mechanism, in order to weight all the detection vectors within a single indicator (D6).

V.4.2 Model

The concept of CSA remains relatively unknown to the cyber community, especially within the French-speaking community. Without necessarily going into the details and the subtleties of the principle, its teaching could be considered within the framework of certain trainings. Indeed, the intellectual approach that it requires to be carried out and its progressiveness seem relevant to any cyber defense analyst (M1).

The abstraction levels of the MCSA and the horizontal or vertical use of knowledge are probably among the most important items to remember for the future: cybersecurity will always be badly perceived by operators and authorities if it diverges from the traditional functioning of the organization by its processes or its dedicated information systems: it must learn to better integrate into operations and share, with the required abstraction, its analysis as close as possible to operators and decision-makers (M2).

The work related to the causality assessment, situation projection and, above all, situation resolution will benefit from being continued more finely by putting into perspective on the one hand, the TTP of the attackers and, on the other hand, models like that of MITRE. This would make it possible to envisage the automatic generation of incident response scenarios adapted to the maritime community (Course of Action) and the continuation of work on the perception of the attacker and his intentions (M3, M4).

At the high level, cyber situation sharing may allow the evolution of standards such as ADatP-3¹ for military purposes, to facilitate the dissemination of cyber threat intelligence and semi-automated cyber incidents reporting between ships, and between ships and centers ashore, including between ships of different flags. This point will be particularly relevant for autonomous ship fleets (integrity checks and reciprocal alerts in the event of an attack to ensure the resilience of the fleet and the accomplishment of the mission) (M5).

¹ADatP-3, Allied Data Processing Publication number 3, is a NATO standard for drafting operational messages. It fixes the patterns of formal messages exchanged between the forces of different countries within the framework of a NATO operation.

V.4.3 Architecture

The continuation of the CARMEN project for anomaly detection on positioning, navigation and time networks and the sharing of cyber tactical situations will continue in parallel with a new research project of the Naval Academy, called Holistic APProach of Integrated Navigation Equipment for cyberSecurity at Sea (HAPPINESS). The objective of this last project is to create an on-board waterproof case equipped with various sensors and AIS, GPS and 4G antennas, allowing real-time transmission of AIS and GPS information in the NMEA format of coastal maritime routes evolution towards shore, in order to inject them on the Naval Cyber Range, but also for the production of additional data sets and the capture of these information in radio format. This work could hopefully feed the evolution and adaptations of the AIS and NMEA standards which have become essential to correct, in depth, the vulnerabilities by design of which they are the object. As the data from these two standards are not necessarily reliable, their correlation with other data sources (satellite imagery, for example) would need to be deepened to increase the confidence level (A1, A2).

On the basis of the work carried out, the information loop from detection to the operator on board the ship seems feasible. It therefore seems relevant to continue work on the autonomous response capacities of the sensors (for example: automatic switch to an integrated system in the event of detection) and on the adaptation of the architecture to allow the fusion of SAs. The use of sensors for the automation of Cyber Battle Damage Assessment offers interesting perspectives in terms of damage assessment and resilience research (A3, A4).

Finally, some of the constraints mentioned for the maritime world are very similar to those encountered in the context of autonomous vehicles, drones and the Internet of Things. The opening of a final perspective towards the application of this work in these constrained fields therefore appears relevant to us (A5).

Summary

In a globalized economy, the maritime sector is essential for the correct functioning of all countries' economies. It plays a key role in the development of countries with maritime borders. Throughout history, the sea has always helped gathering people and developing the economy. The wellness it created has long ago arisen keen interest and, sometimes, engendered conflicts for maritime supremacy. With its ongoing - but often unrecognized - high development, the worldwide shipping industry draws 90% of global world trade. In a highly digitized context, recent studies still show that the cyber maturity level of the sector remains low and behind other essential sectors. Amongst other reasons, this delay could be explained by a difficult perception of the true cyber security level of the sector. While other industries globally managed to achieve this goal, what are the reasons explaining this delay in the maritime sector?

Through an end-to-end analysis, this thesis aims at describing the characteristics of maritime information systems. We list and analyse the specific constraints and working modes of maritime information systems. Then, we define the required processes to build the *Maritime Cyber Situational Awareness* (MCSA). We describe an architecture to achieve MCSA elaboration. This solution, which has been proven both on platform and in an operational context, takes into account the full requirements. Our work then analyses the particularities of both maritime and military worlds to streamline the collected data.

The MCSA elaboration for the maritime sector now seems to be a legitimate and realistic ambition. If it is correctly explained and managed by cyber-security and maritime companies, shipowners, shipyards and public players, this legitimate ambition would represent a major objective for the coming years.

Beyond this conclusion, the analysis and architectures of this study could also be opened and applied to other sectors, such as autonomous vehicles and the Internet of Things (IoT):

it is not unrealistic to think of applying the field of this study to those sectors.

Publications list

1. O. Jacq, X. Boudvin, D. Brosset, Y. Kermarrec, J. Simonin, *Detecting and hunting cyberthreats in a maritime environment: specification and experimentation of a maritime cybersecurity operations centre*, 2018 2nd Cyber Security in Networking Conference (CSNet), 1-8
2. O. Jacq, D. Brosset, J. Simonin, Y. Kermarrec, *Use of Suricata, ElasticStack, Neo4j and Linkurious for network defence*, Suricon 2019
3. O. Jacq, D. Brosset, Y. Kermarrec, J. Simonin, *Cyber attacks real time detection: towards a Cyber Situational Awareness for naval systems*, 2019 International Conference on Cyber Situational Awareness, Data Analytics
4. O. Jacq, PM Laso, D. Brosset, J. Simonin, Y. Kermarrec, MA Giraud, *Maritime Cyber Situational Awareness Elaboration for Unmanned Vehicles*, Maritime Situational Awareness Workshop, 2019

Bibliography

- [A⁺02] National Marine Electronics Association et al. *NMEA 0183–Standard for interfacing marine electronic devices*. NMEA, 2002.
- [ABPG20] Andrej Androjna, Tanja Brcko, Ivica Pavic, and Harm Greidanus. Assessing cyber challenges of maritime navigation. *Journal of Marine Science and Engineering*, 8(10):776, 2020.
- [Al-16] Al-Mohannadi Hamad and Mirza, Qublai and Namanya, Anitta and Awan, Irfan and Cullen, Andrea and Disso, Jules. Cyber-attack modeling analysis techniques: An overview. In *2016 IEEE 4th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW)*, pages 69–76. IEEE, 2016.
- [All] International Maritime Electronics Alliance. An IP/Ethernet Interface Standard for Marine Electronic Devices. <https://www.nmea.org/Assets/onenet%202017%20update.pdf>. Consulté le 20/05/2018.
- [Bar10] Barford, Paul and Dacier, Marc and Dietterich, Thomas G and Fredrikson, Matt and Giffin, Jon and Jajodia, Sushil and Jha, Somesh and Li, Jason and Liu, Peng and Ning, Peng and others. Cyber sa: Situational awareness for cyber defense. In *Cyber situational awareness*, pages 3–13. Springer, 2010.
- [Bea17] Beaumont, Peter and Wolthusen, S. Cyber-risks in maritime container ports: An analysis of threats and simulation of impacts. *ISG MSc Information Security thesis series 2017*, 2017.
- [Ben18] Bruno Bender. Internal report "cybersécurité & domaine maritime", 2018.

- [BG17] Ive Botunac and Marijan Gržan. Analysis of software threats to the automatic identification system. *Brodogradnja: Teorija i praksa brodogradnje i pomorske tehnike*, 68(1):97–105, 2017.
- [Bla18] Félix Blanc. Géopolitique des câbles : une vision sous-marine de l’internet. *Les carnets du centre d’analyse, de prévision et de stratégie*, 2018.
- [Bol19] Bolbot, Victor and Theotokatos, Gerasimos and Boulougouris, Evangelos and Vassalos, Dracos. Safety related cyber-attacks identification and assessment for autonomous inland ships. In *International Seminar on Safety and Security of Autonomous Vessels (ISSAV)*, 2019.
- [BPW14] Marco Balduzzi, Alessandro Pasta, and Kyle Wilhoit. A security evaluation of ais automated identification system. In *Proceedings of the 30th annual computer security applications conference*, pages 436–445, 2014.
- [Bra11] Branlat, Matthieu and Morison, Alexander and Woods, DD. Challenges in managing uncertainty during cyber events: Lessons from the staged-world study of a large-scale adversarial cyber security exercise. In *Human Systems Integration Symposium*, pages 10–25, 2011.
- [Bri19] Jean-Jacques Bridey. La France, puissance maritime. *Revue Défense Nationale*, (8):7–11, 2019.
- [BSG⁺20] Paul M Berges, Basavesh Ammanaghatta Shivakumar, Timothy Graziano, Ryan Gerdes, and Z Berkay Celik. On the feasibility of exploiting traffic collision avoidance system vulnerabilities. *arXiv preprint arXiv:2006.14679*, 2020.
- [Bur14] Ryan Burton. Cybersécurité et marétique, un enjeu européen ? *Centre d’études stratégiques de la marine*, 2014. Consulté le 20/05/2018.
- [Bur16] Joe Burton. Cyber attacks and maritime situational awareness evidence from Japan and Taiwan. In *2016 International Conference on Cyber Situational Awareness, Data analytics and Assessment (CyberSA)*, pages 1–4. IEEE, 2016.
- [BUTN18] Hayretdin Bahşi, Chibuzor Joseph Udokwu, Unal Tatar, and Alexander Nort. Impact assessment of cyber actions on missions or business processes: A systematic literature review. In *ICCWS 2018 13th International Conference on Cyber Warfare and Security*, page 11. Academic Conferences and publishing limited, 2018.

- [Car13] Cardenas, Alvaro A and Manadhata, Pratyusa K and Rajan, Sreeranga P. Big data analytics for security. *IEEE Security & Privacy*, 11(6):74–76, 2013.
- [Cen] Royal Caribbean Press Center. Symphony of the seas fact sheet. <https://www.royalcaribbeanpresscenter.com/fact-sheet/31/symphony-of-the-seas/>. Consulté le 20/05/2018.
- [Chi14] Chifflier, Pierre and Fontaine, Arnaud. Architecture système sécurisée de sonde IDS réseau. *Computer & Electronics Security Applications Rendez-vous (C&ESAR)*, 2014.
- [CIC] CICDE. Glossaire interarmées de terminologie opérationnelle (GIATO).
- [Cim11] Cimpean, Dan and Meire, Johan and Bouckaert, Vincent and Vande Casteele, Stijn and Pelle, Aurore and Hellebooge, Luc. Analysis of cyber security aspects in the maritime sector. Technical report, ENISA, 2011.
- [Cla16] Bryan Clark. Undersea cables and the future of submarine competition. *Bulletin of the Atomic Scientists*, 72(4):234–237, 2016.
- [Com12] North Atlantic Military Committee. Mc 0422/4 NATO military policy on information operations, July 2012.
- [Cos18] Benjamin Coste. *Détection contextuelle de cyberattaques par gestion de confiance à bord d’un navire*. Thèse de doctorat, 2018.
- [CTSdG12] Gabriel D Cavalcante, Sebastien Tricaud, Cleber P Souza, and Paulo Licio de Geus. Interactive analysis of computer scenarios through parallel coordinates graphics. In *International Conference on Computational Science and Its Applications*, pages 314–325. Springer, 2012.
- [da19] Ministère des armées. Droit international appliqué aux opérations dans le cyberspace. *Droit international appliqué aux opérations dans le cyberspace*, 2019. Consulté le 11/09/2019.
- [Das14] Das, Niva and Sarkar, Tanmoy. Survey on host and network based intrusion detection system. *International Journal of Advanced Networking and Applications*, 6(2):2266, 2014.
- [Dia14] Diallo, David and Feuillet, Mathieu. Détection d’intrusion dans les systèmes industriels: Suricata et le cas de Modbus. *C&ESAR2014.(cf. p 44)*, 2014.

- [DiR15] DiRenzo, Joseph and Goward, Dana A and Roberts, Fred S. The little-known challenge of maritime cyber security. In *Information, Intelligence, Systems and Applications (IISA), 2015 6th International Conference on*, pages 1–5. IEEE, 2015.
- [dllea13] Direction de l’information légale et administrative. Livre blanc défense et sécurité nationale, 2013, 2013.
- [dS16] Compagnie Européenne d’Intelligence Stratégique. Impacts sur l’économie française de la fermeture d’un ou plusieurs détroits maritimes majeurs. *Étude prospective et stratégique*, 2016.
- [Dut13] Dutt, Varun and Ahn, Young-Suk and Gonzalez, Cleotilde. Cyber situation awareness: modeling detection of cyber attacks with instance-based learning theory. *Human Factors*, 55(3):605–618, 2013.
- [End95] Mica R Endsley. Toward a theory of situation awareness in dynamic systems. In *Toward a theory of situation awareness in dynamic systems*, volume 37, pages 32–64. SAGE Publications Sage CA: Los Angeles, CA, 1995.
- [Fer97] Jacques Ferber. Les systèmes multi-agents: un aperçu général. *Techniques et sciences informatiques*, 16(8), 1997.
- [Fit15] Fitton, Oliver and Prince, Daniel and Germond, Basil and Lacy, Mark. The future of maritime cyber security, 2015.
- [For13] Ford, Gary and McMahon, Chris and Rowley, Chris. Naval surface ship in-service information exploitation. *Procedia CIRP*, 11:92–98, 2013.
- [Foy05] Foyle, David C and Andre, Anthony D and Hooey, Becky L. Situation awareness in an augmented reality cockpit: Design, viewpoints and cognitive glue. In *Proceedings of the 11th International Conference on Human Computer Interaction*, volume 1, pages 3–9, 2005.
- [Fra14] Franke, Ulrik and Brynielsson, Joel. Cyber situational awareness—a systematic review of the literature. *Computers & Security*, 46:18–31, 2014.
- [Gar18] Phani Kumar Garimella. It-ot integration challenges in utilities. In *2018 IEEE 3rd International Conference on Computing, Communication and Security (ICCCS)*, pages 199–204. IEEE, 2018.

- [GGDM⁺18] G Gonzalez-Granadillo, Samuel Dubus, Alexander Motzek, Joaquin Garcia-Alfaro, Ender Alvarez, Matteo Merialdo, Serge Papillon, and Hervé Debar. Dynamic risk management response system to handle cyber threats. *Future Generation Computer Systems*, 83:535–552, 2018.
- [Goe15] Goetz, Pierre and Cahuzac-Soave, Olivia. Impact de la numérisation sur l'exercice du commandement, 2015.
- [HBB⁺20] Hanan Hindy, David Brosset, Ethan Bayne, Amar Seeam, Christos Tachtatzis, Robert Atkinson, and Xavier Bellekens. A taxonomy of network threats and the effect of current datasets on intrusion detection systems. *IEEE Access*, 2020.
- [Hum15] Christopher Humphries. *User-centred security event visualisation*. Thèse de doctorat, Rennes 1, 2015.
- [Ins] Marine Insight. 10 world's biggest container ships in 2017. <https://www.marineinsight.com/know-more/10-worlds-biggest-container-ships-2017/>. Consulté le 20/05/2018.
- [Jac] Olivier Jacq. Les incidents (connus)... <https://cybermaretique.fr/les-incidents-connus/>. Consulté le 11/09/2019.
- [Jac18] Jacq, Olivier and Boudvin, Xavier and Brosset, David and Kermarrec, Yvon and Simonin, Jacques. Detecting and hunting cyberthreats in a maritime environment: Specification and experimentation of a maritime cybersecurity operations centre. In *2018 2nd Cyber Security in Networking Conference (CSNet)*, pages 1–8. IEEE, 2018.
- [Jac19a] Jacq, Olivier and Brosset, David and Kermarrec, Yvon and Simonin, Jacques. Cyber attacks real time detection: towards a cyber situational awareness for naval systems. In *2019 International Conference on Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*, pages 1–2. IEEE, 2019.
- [Jac19b] Jacq, Olivier and Laso, Pedro Merino and Brosset, David and Simonin, Jacques and Kermarrec, Yvon and Giraud, Marie-Annick. Maritime cyber situational awareness elaboration for unmanned vehicles. In *Maritime Situational Awareness Workshop*, 2019.

- [Jgu16] Ines Jguirim. *Modélisation et génération d'itinéraires contextuels d'activités urbaines dans la ville*. Thèse de doctorat, Ecole doctorale des sciences de la mer, 2016.
- [Jon16] Jones, Kevin D and Tam, Kimberly and Papadaki, Maria. Threats and impacts in maritime cyber security. *Engineering & Technology Reference*, 2016.
- [Kat17] Sokratis K Katsikas. Cyber security of the autonomous ship. In *Proceedings of the 3rd ACM workshop on cyber-physical system security*, pages 55–56. ACM, 2017.
- [Kav18] Kavallieratos, Georgios and Katsikas, Sokratis and Gkioulos, Vasileios. Cyber-attacks against the autonomous ship. In *Computer Security*, pages 20–36. Springer, 2018.
- [KDC⁺18] Alexandre Kabil, Thierry Duval, Nora Cuppens, Gérard Le Comte, Yoran Halgand, and Christophe Ponchel. From cyber security activities to collaborative virtual environments practices through the 3D cybercop platform. In *International Conference on Information Systems Security*, pages 272–287. Springer, 2018.
- [KHL⁺21] Mathias Karlsson, Sandra Haraldson, Mikael Lind, Eddie Olsson, Trond Andersen, and Miluše Tichavska. *Data Visualisation Tools for Enhanced Situational Awareness in Maritime Operations*, pages 355–372. Springer International Publishing, Cham, 2021.
- [Kri13] Krile, Srećko and Kezić, Danko and Dimc, Franc. NMEA communication standard for shipboard data architecture. *NAŠE MORE: znanstveno-stručni časopis za more i pomorstvo*, 60(3-4):68–81, 2013.
- [KSB⁺19] Faris Bugra Kokulu, Ananta Soneji, Tiffany Bao, Yan Shoshitaishvili, Ziming Zhao, Adam Doupé, and Gail-Joon Ahn. Matched and mismatched SOC: A qualitative study on security operations center issues. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, pages 1955–1970, 2019.
- [Kuc07] Kuchar, JE and Drumm, Ann C. The traffic alert and collision avoidance system. *Lincoln Laboratory Journal*, 16(2):277, 2007.

- [Lab20] Richard Labévière. *Reconquérir par la mer : La France face à la nouvelle géopolitique des océans*. Temporis, Paris, 2020.
- [Lag10] Philippe Lagadec. Visualisation et analyse de risque dynamique pour la cyberdéfense. In *Symposium sur la Sécurité des Technologies de l'Information et de la Communication*, pages 3–31, 2010.
- [Lev17] Oskar Levander. Autonomous ships on the high seas. *IEEE spectrum*, 54(2):26–31, 2017.
- [LMWW21] Mikael Lind, Michalis Michaelides, Robert Ward, and Richard Thomas Watson, editors. *Maritime Informatics*. Progress in IS. Springer International Publishing, 2021.
- [LPV13] David López, Oscar Pastor, and L Garcia Villalba. Dynamic risk assessment in information systems: state-of-the-art. In *Proceedings of the 6th International Conference on Information Technology, Amman*, pages 8–10, 2013.
- [Mav17] Mavroeidis, Vasileios and Bromander, Siri. Cyber threat intelligence model: an evaluation of taxonomies, sharing standards, and ontologies within cyber threat intelligence. In *2017 European Intelligence and Security Informatics Conference (EISIC)*, pages 91–98. IEEE, 2017.
- [McG00] McGuinness, Barry and Foy, Louise. A subjective measure of sa: the crew awareness rating scale (cars). In *Proceedings of the first human performance, situation awareness, and automation conference, Savannah, Georgia*, volume 16, 2000.
- [Mer15] Merk, Olaf and Busquet, Bénédicte and Aronietis, Raimonds. The impact of mega-ships, case-specific policy analysis, 2015.
- [mf] Cluster maritime français. Livre bleu de la marétique. <https://issuu.com/opteam/docs/seagital-livre-bleu-12112013>. Consulté le 20/05/2018.
- [Mil19] Natalia Miloslavskaya. *Network Security Intelligence Centres for Information Security Incident Management*. Thèse de doctorat, University of Plymouth, 2019.
- [MK19] Mary C Kay Michel and Michael C King. The future of cyber analytics: Identity classification for systematic and predictive insight. In *2019 International*

- Conference on Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*, pages 1–4. IEEE, 2019.
- [ML17] Pedro Merino Laso. *Détection de dysfonctionnements et d’actes malveillants basée sur des modèles de qualité de données multi-capteurs*. Thèse de doctorat, Ecole nationale supérieure Mines-Télécom Atlantique Bretagne Pays de la Loire, 2017.
- [nat] Marine nationale. Dossier d’information marine 2019. <https://cols-bleus-fr.s3.amazonaws.com/exemplaires/pdf/DIM-2019.pdf>.
- [Nic19] Hellesen Niclas. Cyber situational security awareness architecture (cssa) for industrial control systems. Master’s thesis, NTNU, 2019.
- [Onw12] Onwubiko, Cyril and Owens, Thomas. Review of situational awareness for computer network defense. In *Situational Awareness in Computer Network Defense: Principles, Methods and Applications*, pages 1–9. IGI Global, 2012.
- [oS] Internal Chamber of Shipping. ICS - key facts. <http://www.ics-shipping.org/shipping-facts/key-facts>. Consulté le 20/05/2018.
- [OTA17] OTAN. AAP-06 glossaire OTAN de termes et définitions (anglais et français), 2017.
- [Pan18] Panero, Pablo and Vlsan, Liviu and Brillault, Vincent and Schuszter, Ioan Cristian. Building a large scale intrusion detection system using big data technologies. *PoS*, page 014, 2018.
- [Por14] Porathe, Thomas and Prison, Johannes and Man, Yemao. Situation awareness in remote control centres for unmanned ships. In *Proceedings of Human Factors in Ship Design & Operation, 26-27 February 2014, London, UK*, page 93, 2014.
- [PPM18] Nikolaos Polatidis, Michalis Pavlidis, and Haralambos Mouratidis. Cyber-attack path discovery in a dynamic supply chain maritime risk management system. *Computer Standards & Interfaces*, 56:74–82, 2018.
- [Rø14] Rødseth, Ørnulf Jan and Tjora, Åsmund. A system architecture for an unmanned ship. In *Proceedings of the 13th International Conference on Computer and IT Applications in the Maritime Industries (COMPIT)*, 2014.

- [SHJ⁺19] Shamika N Sirimanne, J Hoffman, W Juan, R Asariotis, M Assaf, G Ayala, H Benamara, D Chantrel, J Hoffmann, A Premti, et al. Review of maritime transport, 2019. Technical report, tech. rep, 2019.
- [Sic18] Franck Sicard. *Prise en compte des risques de cyber-attaques dans le domaine de la sécurité des systèmes cyber-physiques: proposition de mécanismes de détection à base de modèles comportementaux*. Thèse de doctorat, Grenoble Alpes, 2018.
- [Sto11] Stouffer, Keith A and Falco, Joseph A and Scarfone, Karen A. SP 800-82. guide to Industrial Control Systems (ICS) security: Supervisory Control And Data Acquisition (SCADA) systems, Distributed Control Systems (DCS), and other control system configurations such as Programmable Logic Controllers (PLC), 2011.
- [Str17] Strom, Blake E and Battaglia, Joseph A and Kemmerer, Michael S and Kuper-sanin, William and Miller, Douglas P and Wampler, Craig and Whitley, Sean M and Wolf, Ross D. Finding cyber threats with ATT&CK-based analytics. Technical report, Technical Report MTR170202, MITRE, 2017.
- [Svi19] Svilicic, Boris and Kamahara, Junzo and Rooks, Matthew and Yano, Yoshiji. Maritime cyber risk management: An experimental ship assessment. *The Journal of Navigation*, 72(5):1108–1120, 2019.
- [Tad10] Tadda, George P and Salerno, John S. Overview of cyber situation awareness. In *Cyber situational awareness*, pages 15–35. Springer, 2010.
- [Tam18a] Tam, Kimberly and Jones, Kevin. Cyber-risk assessment for autonomous ships. In *2018 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)*, pages 1–8. IEEE, 2018.
- [Tam18b] Tam, Kimberly and Jones, Kevin D. Maritime cybersecurity policy: the scope and impact of evolving technology on international shipping. *Journal of Cyber Policy*, 3(2):147–164, 2018.
- [Tam19] Tam, Kimberly and Jones, Kevin. Factors affecting cyber risk in maritime. In *2019 International Conference on Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*, pages 1–8. IEEE, 2019.

- [TO21] Dominic Thomas and Steven O'Malley. *The Necessity of Standards for Maritime Informatics in Ship Operations*, pages 33–45. Springer International Publishing, Cham, 2021.
- [TXLJ19] Yaodong Tao, Wei Xu, Hongbin Li, and Shenglong Ji. Experience and lessons in building an ics security testbed. In *2019 1st International Conference on Industrial Artificial Intelligence (IAI)*, pages 1–6. IEEE, 2019.
- [Vei12] Veillard, Luc and Anquez, Mathieu and Histrimont, Jean-Pierre. Vulnérabilités de la France face aux flux maritimes. *Étude prospective et stratégique*, page 140, 2012.
- [WDWI16] Cynthia Wagner, Alexandre Dulaunoy, Gérard Wagener, and Andras Iklody. Misp: The design and implementation of a collaborative threat intelligence sharing platform. In *Proceedings of the 2016 ACM on Workshop on Information Sharing and Collaborative Security*, pages 49–56. ACM, 2016.
- [Yad15] Yadav, Tarun and Rao, Arvind Mallari. Technical aspects of cyber kill chain. In *International Symposium on Security in Computing and Communication*, pages 438–452. Springer, 2015.
- [Yen10] Yen, John and McNeese, Michael and Mullen, Tracy and Hall, David and Fan, Xiaocong and Liu, Peng. RPD-based hypothesis reasoning for cyber situation awareness. In *Cyber situational awareness*, pages 39–49. Springer, 2010.
- [Zha19] Zhang, Haiyan and Peng, Minfang and Guerrero, Josep M and Gao, Xingle and Liu, Yanchen. Modelling and vulnerability analysis of cyber-physical power systems based on interdependent networks. *Energies*, 12(18):3439, 2019.
- [Zim14] Carson Zimmerman. Ten strategies of a world-class cybersecurity operations center. *MITRE corporate communications and public affairs. Appendices*, 2014.

Appendices

APPENDIX

A

*Information Technology
and Operational
Technology*

Most commonly noticed differences between Information Technology and Operational Technology are listed in table A.1.

Table A.1: Main differences between Information Technology and Operational Technology. Source: NIST 800.82

Category	IT system	OT system
Performance requirements	Non real-time. Response must be consistent. High throughput is demanded. High delay and jitter may be acceptable. Less critical emergency interaction. Tightly restricted access control can be implemented to the degree necessary for security.	Real-time. Response is time-critical. Modest throughput is acceptable. Access to ICS should be strictly controlled, but should not hamper or interfere with human-machine interaction.
Availability (Reliability) requirements	Responses such as rebooting are acceptable. Availability deficiencies can often be tolerated, depending on the system's operational requirements.	Responses such as rebooting may not be acceptable because of process availability requirements. Availability requirements may necessitate redundant systems. Outages must be planned and scheduled days/weeks in advance. High availability requires exhaustive predeployment testing.

(Table to be followed next page)

Category	IT system	OT system
Risk management requirements	Manage data. Data confidentiality and integrity is paramount. Fault tolerance is less important - momentary downtime is not a major risk. Major risk impact is delay of business operations.	Control physical world. Human safety is paramount, followed by protection of the process. Fault tolerance is essential, even momentary downtime may not be acceptable. Major risk impacts are regulatory non-compliance, environmental impacts, loss of life, equipment, or production.
System operations	Systems are designed for use with typical operating systems. Upgrades are straightforward with the availability of automated deployment tools.	Differing and possibly proprietary operating systems, often without security capabilities built in. Software changes must be carefully made, usually by software vendors, because of the specialized control algorithms and perhaps modified hardware and software involved.
Resources constraints	Systems are specified with enough resources to support the addition of third-party applications such as security solutions.	Systems are designed to support the intended industrial process and may not have enough memory and computing resources to support the addition of security capabilities.

B

Generic cyber risks pour maritime information systems

Our work made it possible to establish a list of the generic risks which could affect a merchant ship, a warship, or a port infrastructure. A majority of the events identified can also apply to maritime drones and autonomous vessels. The approval of an information system before its commissioning¹ provides that a risk analysis is carried out at the information system level, the scope of which is limited to the functional area of the system and its interconnections. The advantage of a more macroscopic view is to have a risk analysis of the ship as a whole, or even by type of ship, or even for the maritime sector as a whole. Although less fine, it allows a better overview.

Navigation systems damage:

- alteration of the electronic navigation system and cartography, leading to loss of vessel positioning and risk of collision;
- loss of traffic surveillance in a maritime area, resulting in an increased risk of collision;
- GPS jamming or spoofing , modification of the geodetic datum on a target vessel or neighboring vessels, resulting in an increased risk of collision;
- alteration of the detection of the navigation radar, resulting in a loss of understanding of the surface situation (Maritime Situational Awareness);

¹See <https://www.ssi.gouv.fr/guide/lhomologation-de-securite-en-neuf-etapes-simples/> for instance

- alteration of the sounder detection (appearance of a "shallow water" alarm...), which could lead to disorganization of the bridge and bad decisions to be made;
- false AIS tracks generation, violations of the integrity or availability of the protocol, leading to the loss of understanding of the surface situation and inappropriate decision-making;
- break in the availability of the ECDIS system, which could impact the navigation of the vessel.

Communication systems damage:

- cut-off of external radio communications, resulting in total isolation of the vessel;
- cut-off of internal communications (VoIP, ToIP, automatic exchange), resulting in disorganisation of the vessel;
- wear of equipment (*eg* due to increased power), resulting in premature degradation and unavailability;
- alteration of the clocks, leading to synchronization faults, inconsistencies, or even malfunctions of on-board installations.

Platform systems damage:

- global attack on the ship's operating system allowing global management of platform installations, making it impossible to act on these installations;
- damage to the electricity sector: high voltage electrical plant, low voltage distribution, or lighting: these attacks can strongly impact the ship: shutdown of cold storage...;
- damage to the auxiliary domain: black water, fresh water, cold air or water production, draining of black water/bilge water, deterioration of reverse osmosis units, loss of refrigerated water;
- damage to the security domain: loss of detection for the fight against disasters, waterways, fire...;

- damage to the mobility domain: impairment on the main thruster control, steering gear, azimuth thruster, dynamic positioning system: erratic start-up, acceleration of the propulsion, blocking of the helm which may cause, in addition to a premature wear, collisions.

Damage to port systems:

- damage to the availability of essential information and functions through the use of ransomware, preventing the port from functioning nominally;
- remote takeover of a installation by exploiting the growing interconnection between systems to perform discrete modification or destructive actions (electricity, pumping...);
- modification of the logistics information systems of the ports (movement of ships, passengers, freight or means of handling and transport), resulting in a disorganization of the port;
- damage in the availability of port services (piloting, refueling, etc.), quays, cranes and storage areas, preventing the reception and disembarkation or departure of ships, etc.
- GPS jamming or spoofing interfering with the automated or semi-automated operation of an unloading gantry crane.

C

Attack modeling on an ECDIS system using MITRE frameworks.

Attack modeling on an ECDIS system is performed using the MITRE PRE-ATT&CK[™] and ATT&CK[™] frameworks. The objective of modeling is also to determine the most appropriate sensors to detect the threat. In case of an advanced and obfuscated threat, in accordance with the recommendations of MITRE, antiviral solutions are deliberately absent from protection tools. The modeling is based on the first tactical and technical phase of reconnaissance and instrumentation, then on a second phase of operation strictly speaking.

The attack method successively uses social engineering methods, in order to send a forged email, writing as the supplier of nautical charts, in order to perform an update on an ECDIS system of a ship at sea, located by its AIS position. The attached update in fact contains a viral charge which performs the encryption of the ECDIS and scans the network in search of other ECDIS in order to perform the encryption.

Although the NIST *framework* is not process-centric, they are best identified here to facilitate understanding of the model. At each step, the cyber sensor which would best allow the identification of the threat is specified (HIDS, NIDS). It is considered that all possible logs are collected by a HIDS.

MITRE focuses its monitoring recommendations on hosts, in particular through the use of *sysmon*, in order to allow the detection of advanced threats not detectable by antiviruses and generally using encrypted network flows. However, the analysis of the types of sensors to be used shows that, in addition to the analysis of logs and activities reported by hosts, anomaly detection by HIDS-type sensors makes it possible to detect violations of file integrity,

for example. Finally, a NIDS type sensor is effective on unencrypted traffic but can also make it possible to identify illegitimate network flows or an anomaly detection based on changes in the volume or behavior of an encrypted network.

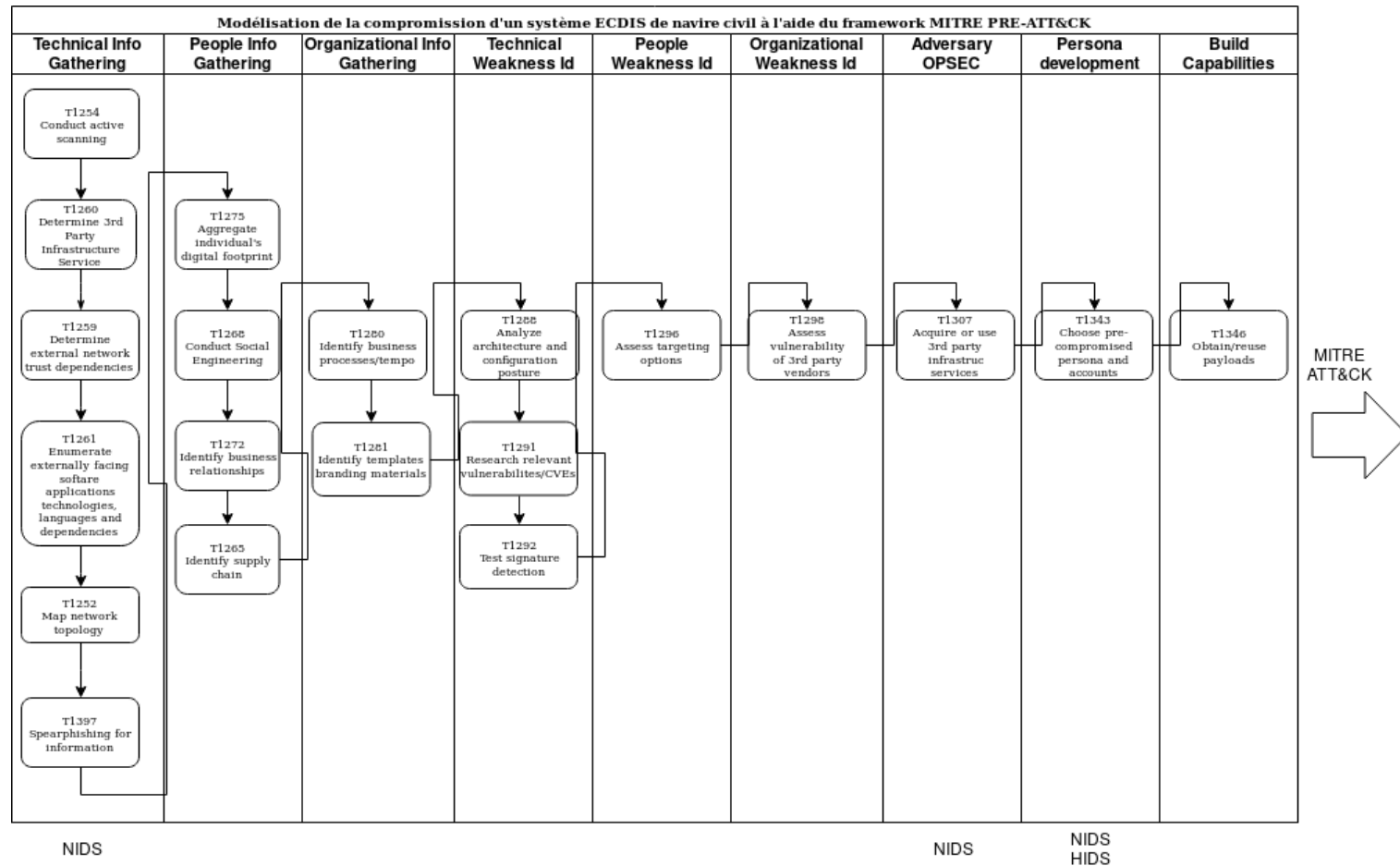


Figure C.1: Pre-exploitation modeling using the MITRE Pre-ATT&CK framework on an advanced attack on an ECDIS system (source: personal work).

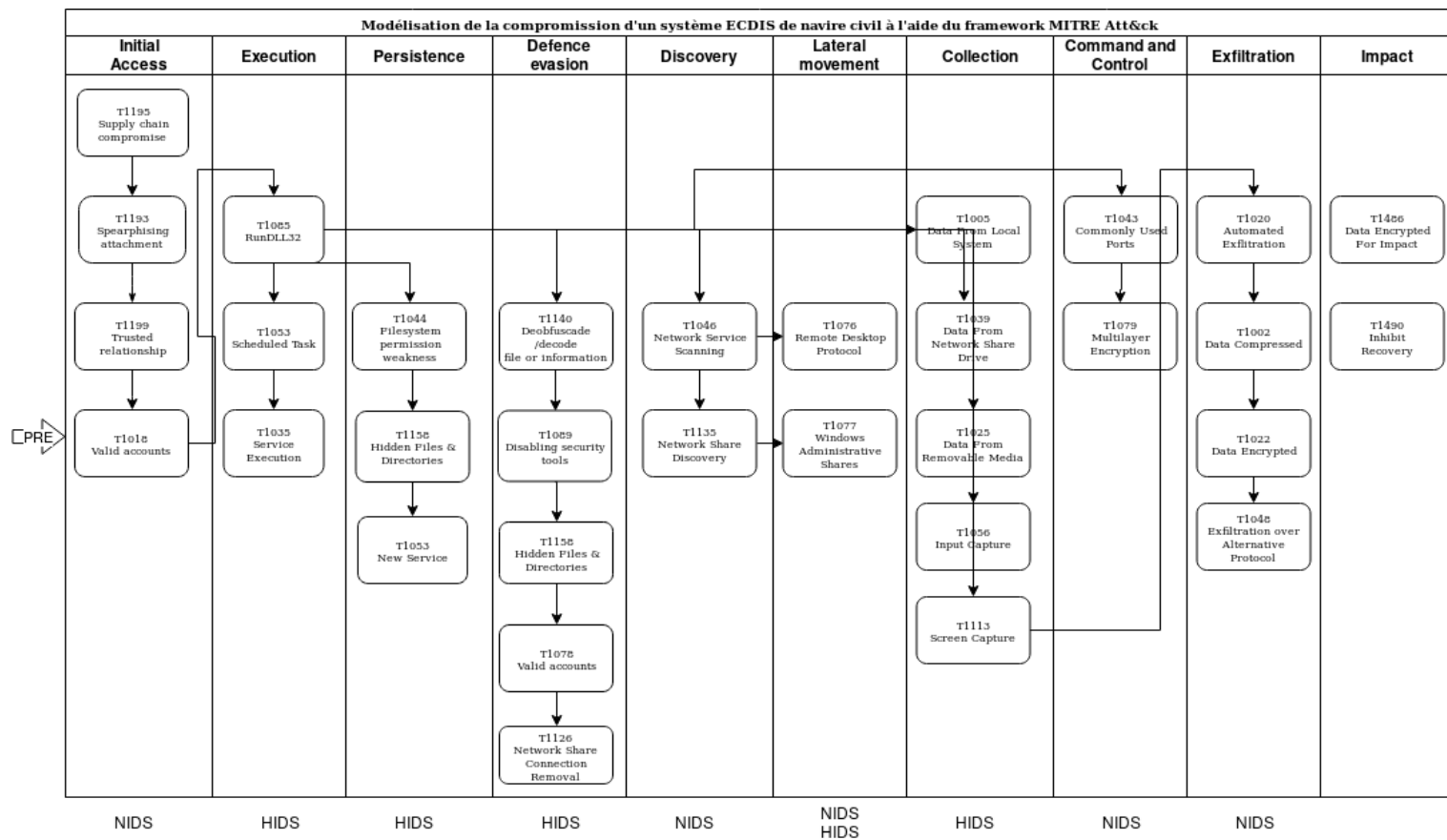


Figure C.2: Post-exploitation modeling using the MITRE ATT&CK framework on an advanced attack on an ECDIS system (source: personal work).

D

Visualisation of a NMEA network determinism based on a talker id-based network traffic capture.

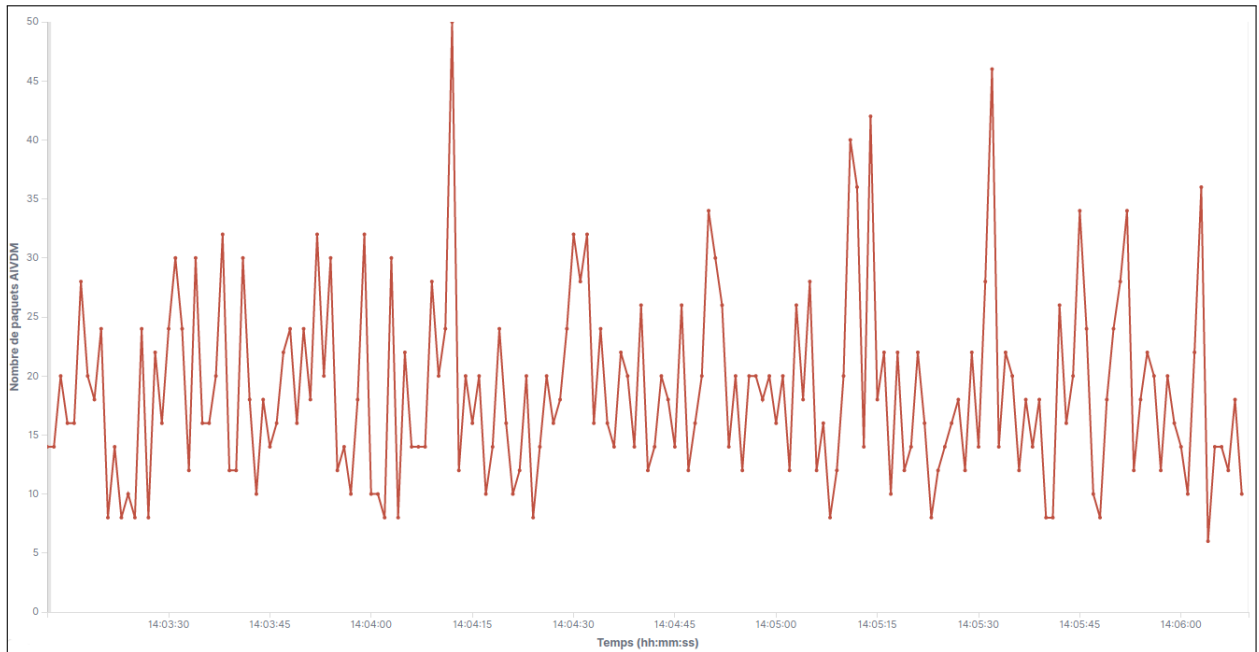


Figure D.1: Evolution over time of the AIVDM sentence (source: personal work).

Figure D.1 shows that the distribution over time of the number of AIVDM sentences does not present, *a priori*, deterministic characteristics. Indeed, the AIS traffic leading to the transmission of AIVDM frames changes over time, depending on the position of the ship, the number of nearby ships, ionospheric propagation, *etc.*



Figure D.2: Evolution over time of the GCGGA sentence (source: personal work).



Figure D.3: Evolution over time of the GCRMC sentence (source: personal work).



Figure D.4: Evolution over time of the GCVTG sentence (source: personal work).



Figure D.5: Evolution over time of the GCZDA sentence (source: personal work).

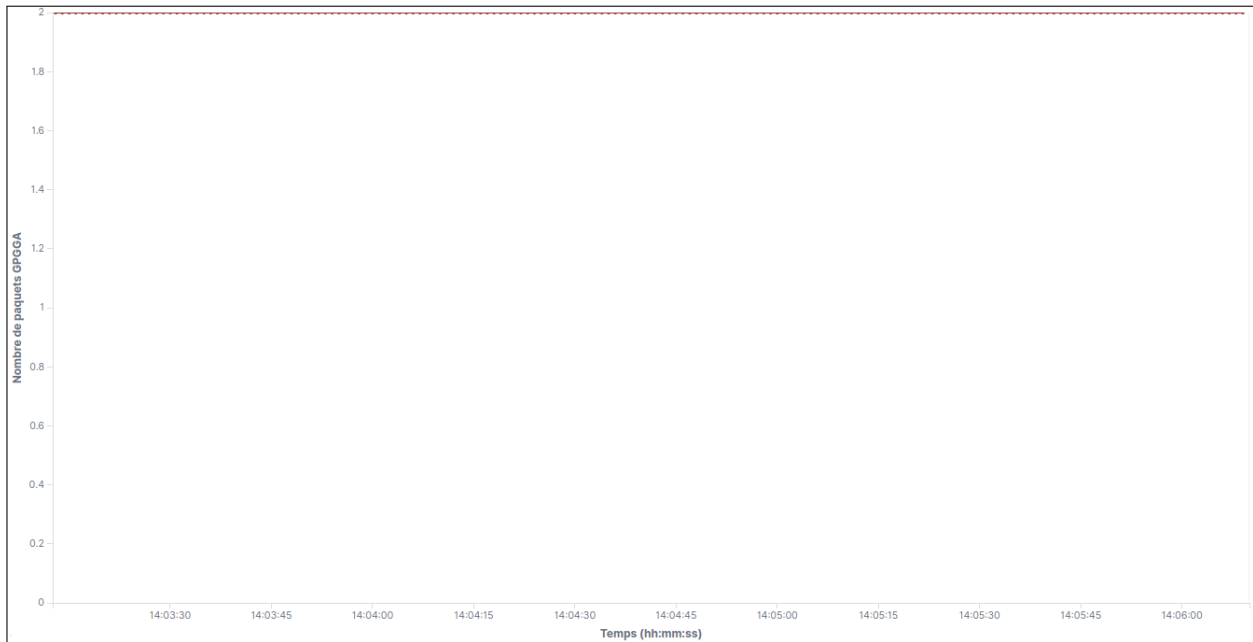


Figure D.6: Evolution over time of the GPGLA sentence (source: personal work).



Figure D.7: Evolution over time of the GPRMC sentence (source: personal work).

All sentences transmitted by the GC talker and shown in the D.2, D.3, D.4 and D.5 figures denote a strong determinism: this GNSS sensor transmits its information on a highly regular and linear basis on the NMEA network: a horizontal line connects the points of emission of the sentences.



Figure D.8: Evolution over time of the IOTXT sentence (source: personal work).

All sentences transmitted by the GP talker and presented in figures D.6 and D.7 denote a strong deterministic character: indeed, this GNSS sensor transmits its information in a very regular and linear basis over the NMEA network. Here too, a horizontal line connects the points of emission of the sentences.

All sentences transmitted by the IO talker and presented in figure D.8 denote a strong deterministic character: indeed, this GNSS sensor transmits its information in a very regular and linear basis over the NMEA network. Here too, a horizontal line connects the points of emission of the sentences.

All sentences transmitted by the NA talker and presented in figures D.9, D.10 and D.11 denote a strong deterministic character: indeed, this sensor transmits its information in a very regular and linear basis over the NMEA network. Here too, a horizontal line connects the points of emission of the sentences.

All sentences transmitted by the NW talker and presented in figures D.12, D.13, D.14 and D.15 denote a strong deterministic character: this sensor transmits its information in a very regular and linear basis over the NMEA network. Here too, a horizontal line connects the points of emission of the sentences.

All sentences transmitted by the PH talker (proprietary format) and presented in figures



Figure D.9: Evolution over time of the NAGGA sentence (source: personal work).



Figure D.10: Evolution over time of the NAVHW sentence (source: personal work).



Figure D.11: Evolution over time of the NAVTG sentence (source: personal work).

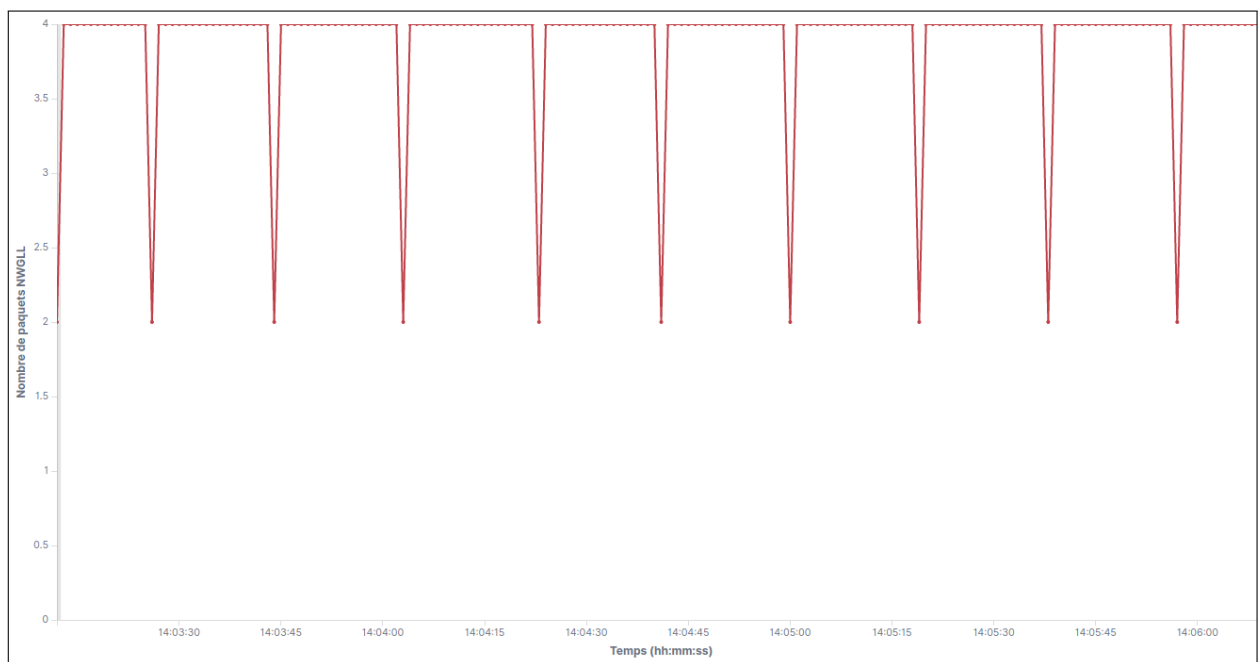


Figure D.12: Evolution over time of the NWGLL sentence (source: personal work).

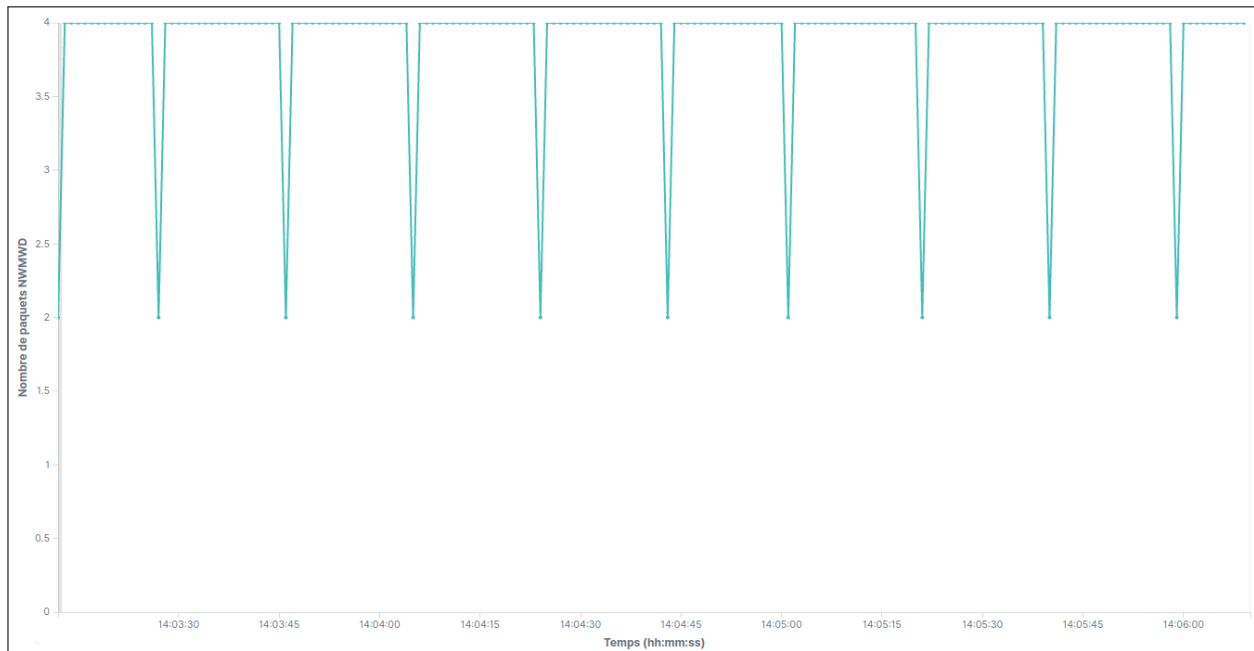


Figure D.13: Evolution over time of the NWMWD sentence (source: personal work).

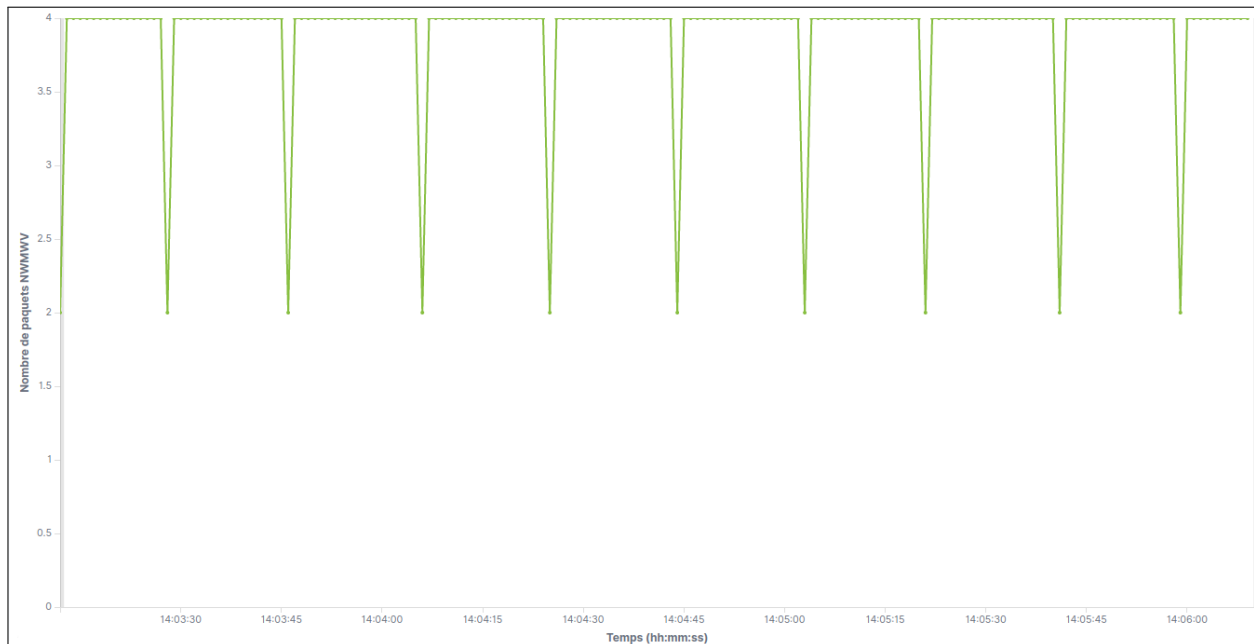


Figure D.14: Evolution over time of the NWMWV sentence (source: personal work).

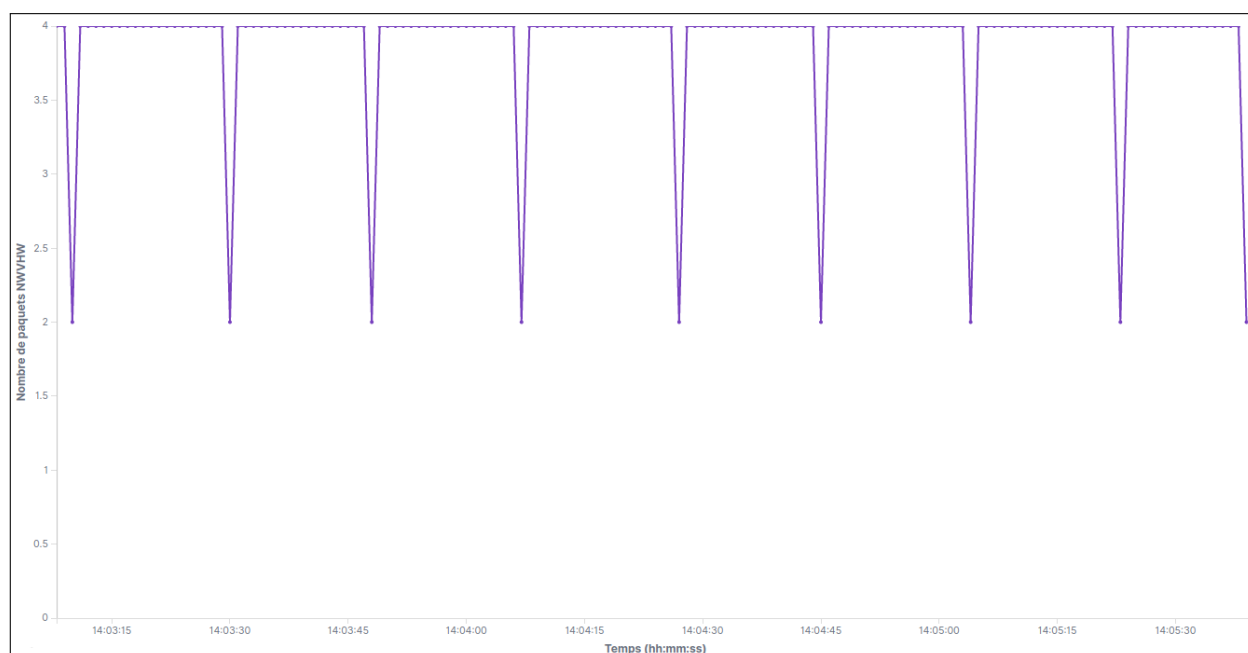


Figure D.15: Evolution over time of the NWVHW sentence (source: personal work).

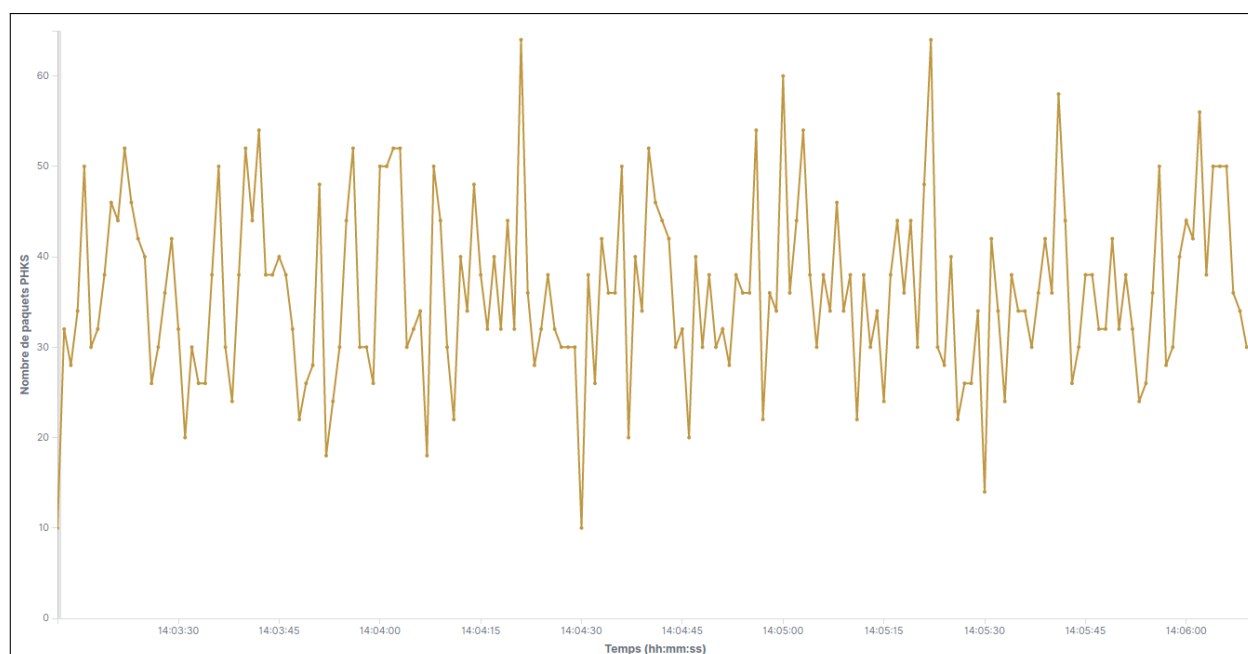


Figure D.16: Evolution over time of the PHKS sentence (source: personal work).

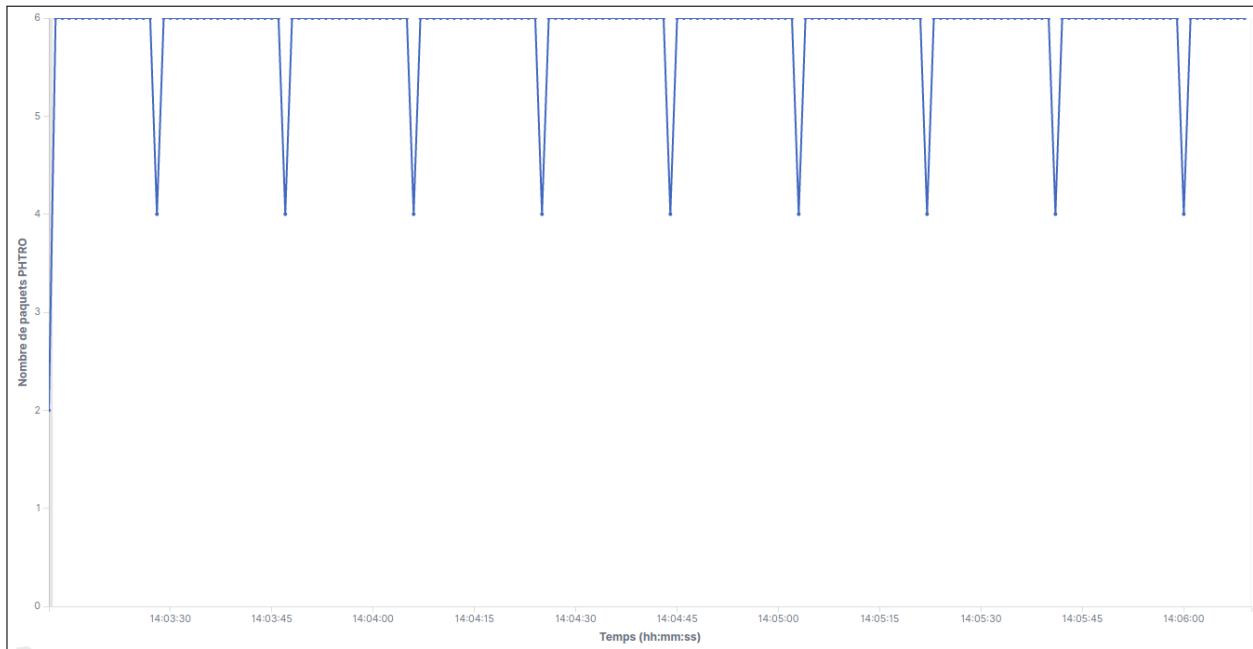


Figure D.17: Evolution over time of the PHTRO sentence (source: personal work).

D.16 and D.17 denote a strong deterministic character for the TRO sentence: this sensor transmits its information in a very regular and linear basis over the network. An interesting comparison can be made with the NW sensor, the repetition pattern being very similar. On the other hand, the PHKS sentence does not present an obvious deterministic character.



Figure D.18: Evolution over time of the RAOSD sentence (source: personal work).



Figure D.19: Evolution over time of the RARSD sentence (source: personal work).

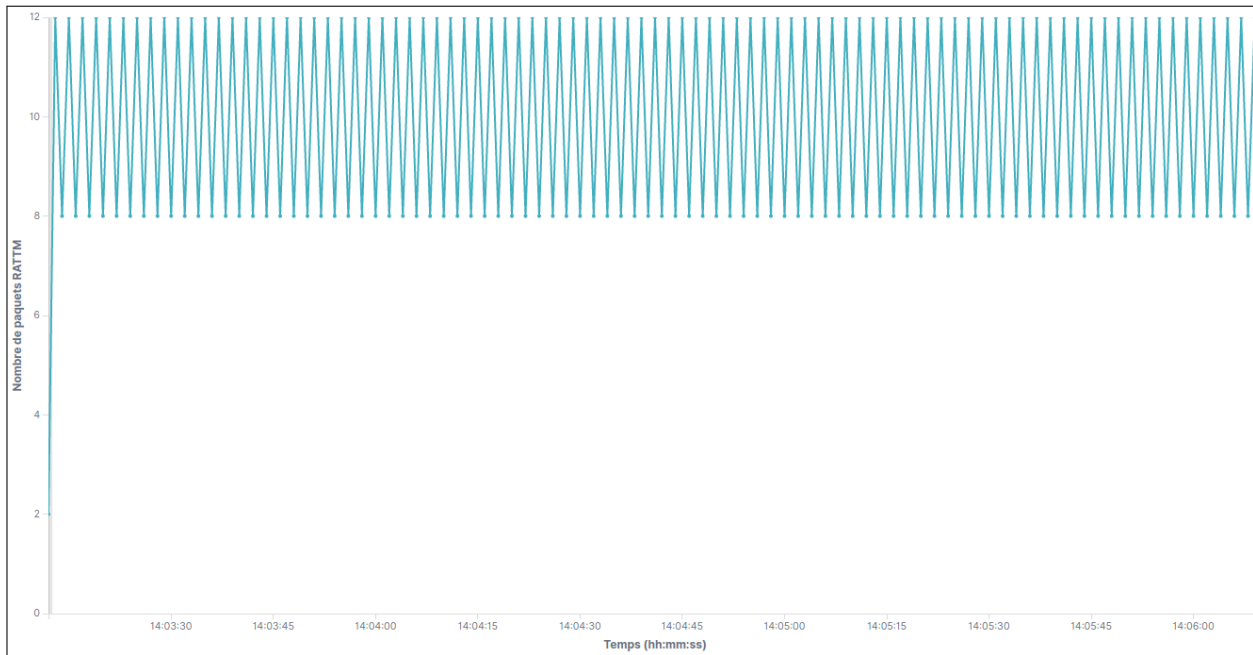


Figure D.20: Evolution over time of the RATTM sentence (source: personal work).

All sentences emitted by the RA talker and presented in figures D.18, D.19 and D.20 denote a strong character deterministic: this radar sensor transmits its information in a very regular and linear manner on the NMEA network. Here too, a horizontal line connects the points of emission of the RAOSD and RARSD sentences. However, the RATTM sentence (D.20) transmits its information in a more specific way: the number of transmitted TTM sentences changes regularly, but at a relatively high frequency.

All sentences emitted by the SD talker and presented in figures D.21 and D.22 denote a strong deterministic character: this sensor transmits its data in a very regular and linear manner on the NMEA network. Here too, a horizontal straight line connects the emission points.

All sentences emitted by the SP talker and presented in figures D.23 and D.24 denote a strong deterministic character: this sensor transmits its data in a very regular and linear manner on the NMEA network. Here too, a horizontal straight line connects the emission points.



Figure D.21: Evolution over time of the SDDBT sentence (source: personal work).



Figure D.22: Evolution over time of the SDDPT sentence (source: personal work).



Figure D.23: Evolution over time of the SPBDT sentence (source: personal work).



Figure D.24: Evolution over time of the SPDPT sentence (source: personal work).

Glossary

ACARS Aircraft Communications and Reporting System. 42

ADS-B Automatic Dependent Surveillance-Broadcast. 42

AIS Automatic Identification System. 20, 43, 44, 69, 100–104, 106, 108, 124, 146, 149, 153

ANSSI Agence Nationale de la Sécurité des Systèmes d'Information. 3, 72

APT Advanced Persistent Threat. 52

ARPA Automatic Radar Plotting Aid. 101, 106

ASCII American Standard Code for Information Interchange. 101

BCP Business Continuity Plan. 60

BM Bandwidth Manager. 74

BNWAS Bridge Navigational Watch Alarm System. 20

CARMEN Cyber Situational Awareness Made Easy for NMEA. 110, 112–114, 118, 120, 124

CCR Cargo Control Room. 20

CCTV Closed Circuit Television. 20

CE Central Engine. 74, 77, 89, 92, 97

CMS Combat Management System. 44

CMS Content Management System. 98

- CNV** Coopération Navale Volontaire. 3
- CoA** Course of Action. 63, 109
- COMEX** Compagnie Maritime d'Expertise. 16
- CP** Central Processor. 74, 77
- CPA** Closest Point of Approach. 44, 45
- CSA** Cyber Situational Awareness. 46–48, 50, 56, 65, 123
- CSAC** Cyber Situational Awareness Console. 73, 75, 77
- CSIRT** Computer Security Incident Response Team. 34, 48, 80, 98, 114
- CTI** Cyber Threat Intelligence. 34, 47, 48, 52, 59, 61, 74
- CVE** Common Vulnerabilities and Exposures. 98, 100
- CVSS** Common Vulnerability Scoring System. 98
- DECT** Digital Enhanced Cordless Telephone. 20
- DFIR** Digital Forensics and Incident Response. 74
- DNS** Domain Name Service. 74, 92
- DP** Dynamic Positioning. 20
- DRA** Dynamic Risk Assessment. 96, 97, 121
- DRP** Disaster Recovery Plan. 60
- ECDIS** Electronic Chart Display Information System. 20, 44, 53, 55, 100, 107, 115, 120, 146, 149
- ECS** Electronic Chart System. 44, 53, 55, 69, 100, 102, 109, 111, 112, 115, 118–120
- EEZ** Economic Exclusive Zone. 15, 16
- EMP** Electro Magnetic Pulse. 30
- ETA** Estimated Time of Arrival. 45

- FN** False Negative. 58
- FP** False Positive. 57
- FPC** Full Packet Capture. 73
- GDP** Gross Domestic Product. 15
- GMDSS** Global Maritime Distress and Safety System. 20
- GNSS** Global Navigation Satellite System. 20, 69, 88, 101, 102, 107–109, 118, 156, 157
- GPS** Global Positioning System. 20, 101, 102, 107, 109, 111–113, 124, 145, 147
- GSM** Global System for Mobile communications. 20
- HAPPINESS** Holistic APProach of Integrated Navigation Equipment for cyberSecurity at Sea. 124
- HIDS** Host-based Intrusion Detection System. 37, 50, 53, 58, 59, 61, 72, 75, 78, 98, 149
- HMI** Human Machine Interface. 119
- HUD** Head Up Display. 41, 44
- IBS** Integrated Bridge System. 58, 102
- ICMP** Internet Control Message Protocol. 80
- ICS** Industrial Control System. 52, 53, 142
- IFREMER** Institut Français de Recherche pour l'Exploitation de la Mer. 16
- IMO** International Maritime Organisation. 18, 19
- INS** Integrated Navigation System. 102, 103, 105, 121
- IoC** Indicator of Compromise. 48, 52, 58, 59, 74, 80
- IP** Internet Protocol. 20
- IT** Information Technology. 19–22, 28, 35, 36, 86, 91, 117
- KT** Knowledge of Them. 40, 42–44, 47, 48, 50, 58, 59, 62, 100

- KU** Knowledge of Us. 40, 42–44, 47, 48, 50, 58, 59, 62, 92, 96, 100, 119
- LE** Local Engine. 73, 87, 89, 120
- LP** Local Preprocessor. 73
- MASS** Maritime Autonomous Surface Ship. 18
- MCSA** Maritime Cyber Situational Awareness. 7, 33, 51, 55, 56, 62, 65, 67, 68, 72, 75, 82, 91, 96, 100, 114, 119, 121, 123
- MFD** Multi-Functions Display. 44
- MIS** Maritime Information Systems. 19, 22–28, 30, 31, 33–35, 51–55, 58, 60, 65, 67, 75, 88, 91, 92, 96, 100, 110
- MRE** Marine Renewable Energies. 16, 19, 21
- MSA** Maritime Situational Awareness. 100, 101
- MTU** Maximum Transmission Unit. 84
- NATO** North Atlantic Treaty Organization. 24, 38
- NAVWAR** Navigation War. 101
- NCS** Network Connexion Safety. 72
- NIDS** Network-based Intrusion Detection System. 37, 50, 53, 58, 59, 61, 62, 72, 78–81, 85, 87, 93, 94, 96, 114, 149, 150
- NIST** National Institute of Standards and Technology. 19, 69
- NMEA** National Marine Electronics Association. 44, 69, 91, 100, 102–106, 110, 111, 114, 118, 121, 124, 156, 157, 164
- NPI** Network Probe Isolation. 72
- NTP** Network Time Protocol. 88
- OECD** Organisation for Economic Co-operation and Development. 12
- OSI** Open Systems Interconnection. 92, 94

- OSINT** Open Source INTelligence. 30
- OT** Operational Technology. 19–22, 28, 35, 36, 53, 69, 85, 86, 91, 110, 114, 117
- PLC** Programmable Logical Controller. 22, 30, 52, 70
- PMS** Property Management System. 20
- RTT** Round Trip Time. 82
- SA** Situational Awareness. 7, 40–46, 48, 55, 56, 61, 76, 77, 115, 121
- SDR** Software Defined Radio. 111
- SGDSN** Secrétariat Général de la Défense et de la Sécurité nationale. 3
- SHOM** Service Hydrographique et Océanographique de la Marine. 16
- SIEM** Security Information and Event Management. 34, 36–38, 48, 62, 78–80, 88, 89, 97, 120
- SIRP** Security Incident Response Platform. 34, 60, 74
- SMTP** Simple Mail Transfer Protocol. 79
- SNLE** Sous-marin Nucléaire Lanceur d’Engins. 16, 17
- SOC** Security Operations Center. 34, 36, 38, 48, 50, 52, 54, 58, 62, 73, 80, 92, 97, 110, 114
- SSAS** Shipboard Security Alarm System. 20
- SSM** Ship Shore Manager. 73, 74, 84, 85
- STIX** Structured Threat Information eXpression. 48
- TAP** Test Access Port. 72
- TAXII** Trusted Automated eXchange of Indicator Information. 48
- TCAS** Traffic Collision Avoidance System. 42, 44
- TCP** Transport Control Protocol. 93, 103
- TN** True Negative. 58

TP True Positive. 58

TTP Tactics, Techniques, Procedures. 34, 48, 61, 63, 109, 123

UDP User Datagram Protocol. 102, 103

USV Unmanned Surface Vehicle. 101

VDR Voyage Data Recorder. 20

VLAN Virtual Local Area Network. 20

VPN Virtual Private Network. 20, 73

VSAT Very Small Aperture Terminal. 82

VTs Vessel Traffic Service. 44

WiFi Wireless Fidelity. 20

Titre : Détection, analyse contextuelle et visualisation de cyberattaques en temps-réel : élaboration de la *Cyber Situational Awareness* du monde maritime

Mots clés : cybersécurité ; maritime ; navire ; port ; détection d'intrusion ; situational awareness

Dans une économie globalisée, le secteur maritime est essentiel au bon fonctionnement des économies et permet d'acheminer 90% des marchandises. Dans un contexte de forte numérisation, le niveau de cybersécurité du secteur maritime reste en retrait par rapport aux autres secteurs d'activité d'importance vitale. Au travers d'une analyse de bout en bout, cette thèse décrit les particularités des systèmes d'information maritimes et modélise le concept de *Maritime Cyber Situational Awareness*.

Ensuite, une proposition d'architecture est décrite pour permettre l'acquisition de cet état de connaissance. Cette solution, éprouvée sur plate-forme, a permis de répondre à l'ensemble des critères d'élaboration. Enfin, les travaux soulignent et détaillent les spécificités du monde maritime pour tirer un profit maximal des données issues de la cybersurveillance. Les analyses et architectures de cette étude dans un contexte contraint pourront probablement être élargies à d'autres secteurs, comme par exemple les véhicules autonomes ou encore l'Internet des objets.

Title: Real-time detection, contextual analysis and visualisation of cyber-attacks: elaboration of the Maritime Cyber Situational Awareness

Keywords: cybersecurity ; maritime ; ship ; port ; intrusion detection ; situational awareness

Abstract: In a globalized economy, the maritime sector plays an essential role for the countries' economies, drawing 90% of the global world trade. In a highly digitalized transformation context, the cybersecurity level of the maritime sector remains low compared to other essential sectors. Through an end-to-end analysis, this thesis aims at describing the unique combined characteristics of maritime information systems. Then, we apply situational awareness definition to maritime cybersecurity and model the concept of Maritime Cyber Situational Awareness.

Then we describe the proposal of an architecture to achieve MCSA elaboration, which has been tested and proven on our experimental platform, taking into account the full requirements. Our work then analyses the particularities of the maritime world to streamline the collected data. The analysis and architectures of this study could also be opened and applied to other sectors, such as autonomous vehicles and the Internet of Things (IoT).